

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 28, 2002, 11:29:28 ; Search time 66.93 seconds
(without alignments)
10470.527 Million cell updates/sec

Title: US-09-509-188-3
Perfect score: 2853
Sequence: 1 ggaacccacacacacacacg.....ccgccttgatcttaact 2853

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA :
1: /cgn2_6/pdata/1/lna/5A.COMB.seq : *
2: /cgn2_6/pdata/1/lna/5B.COMB.seq : *
3: /cgn2_6/pdata/1/lna/6A.COMB.seq : *
4: /cgn2_6/pdata/1/lna/6B.COMB.seq : *
5: /cgn2_6/pdata/1/lna/PTUS.COMB.seq : *
6: /cgn2_6/pdata/1/lna/Backfile1.seq : *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	103.2	3.6	19124	2	US-08-487-826B-13 Sequence 13, Appl
2	99.2	3.5	19124	2	US-08-487-826B-13 Sequence 13, Appl
3	95	3.3	6124	4	US-08-213-419B-3 Sequence 3, Appl
4	86.8	3.0	1511	1	US-07-991-867B-8 Sequence 8, Appl
5	86.8	3.0	1511	1	US-08-107-755A-8 Sequence 8, Appl
6	86.8	3.0	1511	2	US-08-544-332-8 Sequence 8, Appl
7	86.6	3.0	7218	1	US-08-232-463-14 Sequence 14, Appl
8	84.4	3.0	6124	4	US-08-213-419B-3 Sequence 3, Appl
9	83.2	2.9	3440	5	US-08-471-791-27 Sequence 27, Appl
10	83.2	2.9	3440	5	US-08-471-791-27 Sequence 27, Appl
11	82.2	2.9	16442	2	US-08-883-795A-36 Sequence 36, Appl
12	80.6	2.8	665	2	US-08-883-795A-36 Sequence 36, Appl
13	79	2.8	9636	1	US-08-323-170B-1 Sequence 1, Appl
14	79	2.8	9636	1	US-08-323-170B-1 Sequence 1, Appl
15	77.6	2.7	615	4	US-08-954-441-1 Sequence 186, App
16	76.6	2.7	3095	6	US-08-954-441-1 Sequence 186, App
17	75.6	2.7	3095	6	US-08-954-441-1 Sequence 186, App
18	75.4	2.6	1511	1	US-07-991-867B-8 Sequence 8, Appl
19	75.4	2.6	1511	1	US-08-107-755A-8 Sequence 8, Appl
20	75.4	2.6	1511	2	US-08-544-332-8 Sequence 8, Appl
21	75.4	2.6	3489	2	US-08-728-323A-1 Sequence 1, Appl
22	75.4	2.6	32207	2	US-08-728-323A-1 Sequence 1, Appl
23	75.4	2.6	32207	2	US-08-728-323A-1 Sequence 1, Appl
24	75.4	2.6	32207	4	US-08-757-669A-20 Sequence 20, Appl
25	75	2.6	837	4	US-09-230-371A-20 Sequence 20, Appl
26	75	2.6	51259	3	US-08-998-416-288 Sequence 288, App
27	74.4	2.6	1431	4	US-08-781-891-209 Sequence 209, App
					Sequence 2, Appl

28	74	2.6	732	4	US-08-998-416-1036 Sequence 1036, Ap
29	73.2	2.6	731	1	US-08-451-405A-2 Sequence 2, Appl
30	72	2.5	6152	4	US-08-973-463-1 Sequence 1, Appl
31	72	2.5	6768	1	US-08-107-755A-1 Sequence 1, Appl
32	72	2.5	8457	1	US-07-991-867B-1 Sequence 1, Appl
33	72	2.5	8457	1	US-08-544-332-1 Sequence 1, Appl
34	71.8	2.5	1241	1	US-07-593-657-6 Sequence 6, Appl
35	71.8	2.5	1241	4	US-08-942-012B-3 Sequence 3, Appl
36	71.6	2.5	636	4	US-08-998-416-1137 Sequence 1137, Ap
37	71	2.5	636	4	US-08-998-416-1137 Sequence 1137, Ap
38	71	2.5	854	4	US-08-998-416-534 Sequence 534, App
39	70.6	2.5	665	2	US-08-883-795A-36 Sequence 36, Appl
40	70.6	2.5	834	4	US-08-998-416-305 Sequence 305, App
41	70.4	2.5	711	4	US-08-998-416-786 Sequence 786, App
42	70.2	2.5	5852	1	US-07-867-106-2 Sequence 2, Appl
43	70.2	2.5	8920	2	US-08-446-853A-1 Sequence 1, Appl
44	70.2	2.5	8920	4	US-09-150-741-1 Sequence 1, Appl
45	70.2	2.5	12001	1	US-08-458-568A-11 Sequence 11, Appl

ALIGNMENTS

RESULT 1
US-08-487-826B-13
Sequence 13, Application US/08487826B
Patent No. 5993827
GENERAL INFORMATION:
APPLICANT: Slim, Kim L.
APPLICANT: Chitnis, Chetan
APPLICANT: Miller, Louis H.
APPLICANT: Peterson, David S.
APPLICANT: Su, Xin-zhaun
APPLICANT: Wellens, Thomas E.
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobe Martens Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: California
COUNTRY: US
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,826B
FILING DATE: 10-SEP-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Israelson, Ned
REGISTRATION NUMBER: 29,655
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 19124 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-487-826B-13

Query Match

3.6%; Score 103.2; DB 2; Length 19124;

QY 374 gngcatngtcatcagttatgcnctttttttttgancgctgtatgttgcctt 433
 Db 747 taaactatattatgtatcttcttttttttttttttttttttttttttttt 1111
 QY 434 attgagtgctgcnacgtgaagaattgaaacatttattttggg-----gaaatat 488
 Db 807 attatataatcattatggaataaatttcccttttaggaattgaggtcattatga 866
 QY 489 aatttaattgttcaatataatagaagaataattatncccttgatgttactgtatgcg 548
 Db 867 gatttaaaaaaaatataatataatataatataatataatattatgtatgtatga 926
 QY 549 agtgaagaatcttggaaatatttggaaacttgccttttccaaaaagtaaatattg 608
 Db 927 acatatataatttttttaatatatttgaagaagaagaatgaataatgcatattg 986
 QY 609 atagtacttaagttacacatgaaatgaataaaaaaaatgaataatagaataaac 668
 Db 987 taatagctatataactaatattgtatttttcatcttttaatttttagctatataat 1046
 QY 669 tgaatgactacccctcaagtttgaactatttctgttgccttccctcaactcta 728
 Db 1047 atacatatataatgaataatataaaatgattgggtgattatcccaaatatttta 1106
 QY 729 agtccccaacataaataatttcatattgtcatattctatcttttagaagtgaaaca 788
 Db 1107 ttaagaaccttataaataatgtatnaaatgtttataaaaaaatatataatgtt 1166
 QY 789 aaatattatcaagttatataatgttttttcaataaagaataaataaataaata 848
 Db 1167 aagtaactagaataatattttttccatcatatttttccaataatagaataaata 1226
 QY 849 atagtagttacaaaaaaataaata--atatttaccagctgtaaaaaacataaac 906
 Db 1227 ggtatacttataatcatataatgaataatgtattttatnaaatatataatgaac 1286
 QY 907 ctaaacaccttaatttaacttttaggttaaacaccttaacacttttgataaatcttaaac 966
 Db 1287 ataaaatattgaagactttatgtatggaagaacataatattatgaagccttataat 1346
 QY 967 taaacataaactaaaccttaacttaacttaacttaacaccttaagtttttaagtt 1026
 Db 1347 ttgttaataataaataaataaataaataaataaataaagctttagtaataatgaat 1406
 QY 1027 tagtttttttattatagtttaggttttccaaagtttaagtttaccagaagttt 1086
 Db 1407 aatgaatttctttgttgaatattatattatgtatgtatttccgagaccagaataaa 1466
 QY 1087 atggtttgagattatgaacttagattagttttagttttagcagagcttcaagttttt 1146
 Db 1467 aaaaatcaagcttattgttcgataatcaaatgataataaataatatttttttaac 1526
 QY 1147 taaaaaataatttttttgaacactctatttttttttttttttttttttttttttt 1199
 Db 1527 tccatagaatattttaaagaataataatgcgactatataatcttgttattgtaaaaaactt 1586
 QY 1200 ttataataaataaataaataatatttaactcatcgttttcatattgaagtgtcatgta 1259
 Db 1587 ttatattggaagaataaataatatttatattattatttttttttttttttttttt 1646
 QY 1260 acattatttttttggtaaaaaaatgtccacttttagaattcccaatgcaaaaattatta 1319
 Db 1647 ttataatttaatt 1706
 QY 1320 ttttttgaagctaaatgaatgcaagtgatgtactttaaataatatttttttttttt 1379
 Db 1707 tggagatcaaatatttttaatttttttttttttttttttttttttttttttttttt 1766
 QY 1380 aatgcatatttggtaaaatgttgaatataagaacttaattatatttatttttttt 1439
 Db 1767 ataactaatcccttgttatttcttaaaaaaaatgtacacatgttaacataatgtatataat 1826

QY 1440 ttcttaactctgtgta-----aaatgtcaagtaaaatttatttagaagaacttgatga 1494
 Db 1827 atataatataatgtatttttcttaacttaataatatttttagaataatttaatttttt 1886
 QY 1495 atatttgttcaatttttttaaagatatcgatattgtaataaacaattttattgtatga 1554
 Db 1887 gaataactataattttgtaaacgcgaatcttaaaatgttcaaaaaatgagaatgaaaaa 1946
 QY 1555 tgaacctaaatctcatccttaagaaggttgaacgaagaataagtaacagtttgggaaa 1614
 Db 1947 ctccatttataatataatataatataatattatattattattattatttttttttt 2006
 QY 1615 gctactatggcccaagtaactaaatcttttttttttttttttttttttttttttttt 1674
 Db 2007 ttttaataatagtt 2066
 QY 1675 tagttagagttaaatgaacacacatgattacatcatattagaaatttttcaaaaaatt 1734
 Db 2067 attttttttgaaatgaataaataaagaaggtcatattcttgattattttctttatgag 2126
 QY 1735 actaggaataatgaataatcttttttttttttttttttttttttttttttttttttt 1794
 Db 2127 aatgtcaaaaaaaatgaataaataaataaataaataaataaataaataaataaataa 2184
 QY 1795 atatactaaacaaattttaaataaagaataaaggaacatgagatcaataaataat 1854
 Db 2185 ataatataatataatatttattttaaattttatataatataatataatataatatttt 2244
 QY 1855 gttattttcttaagatgattgattatttataatataatccgctcatatatttcaanaata 1914
 Db 2245 aaatataatattgtattgttccaaaaaaataaataaataaataaataaataaataa 2296
 QY 1915 ataatatttttgcgtacgaataatcatatataatcttaataaaccagctgtatgt 1974
 Db 2297 atattataaatacatataattataataacataaagaataaataaataaataaataa 2356
 QY 1975 atattttattatataatataatataaaccataacgccttaacatgataacatcaaa 2034
 Db 2357 aaaaattaaagttcttaaaaattatataatataatataatataatataatataatata 2416
 QY 2035 actttttcttgcgttgcataactcaaggttgcgaagaatgaataaataaataaataa 2092
 Db 2417 atattccttgttttcatattgttgaagattgaaaaaaataaataaataaataaataa 2474

RESULT 4
 US-07-991-867B-8
 Sequence 8, Application US/07991867B
 Patent No. 5476781
 GENERAL INFORMATION:
 APPLICANT: Moyer, Richard W.
 APPLICANT: Hall, Richard L.
 APPLICANT: Gruidl, Michael E.
 TITLE OF INVENTION: No. 5476781e1 Entomopoxvirus Expression System
 NUMBER OF SEQUENCES: 66
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: David R. Saliwanichik
 STREET: 2421 N.W. 41st Street, Suite A-1
 CITY: Gainesville
 STATE: FL
 COUNTRY: USA
 ZIP: 32606
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/991,867B
 FILING DATE: 12-DEC-1992
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: WO 92/14818

FILING DATE: 12-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/827,685
FILING DATE: 30-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/657,584
FILING DATE: 19-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanichik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: UFI114.C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 1511 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Amsacta moorei entemopoxvirus
FEATURE:
NAME/KEY: CDS
LOCATION: complement (18..218)
FEATURE:
NAME/KEY: CDS
LOCATION: complement (234..782)
FEATURE:
NAME/KEY: CDS
LOCATION: 852..1511
US-07-991-867B-8

Query Match 3.0%; Score 86.8; DB 1; Length 1511;
Best Local Similarity 43.8%; Pred. No. 2.2e-07;
Matches 476; Conservative 0; Mismatches 607; Indels 5; Gaps 2;

QY 915 taaatttaacttttagttaaacccttaaccccttgatgaataacttaacatcaatcatt 974
DB 196 TAATTAATTAATCACTTTTTCATTTTGAATTAATAAATAATTAATTAATTAATCAAT 255
QY 975 aaacacctaaccctaactcctaactcctaaccctaagtgcttaagtgtagtgc 1034
DB 256 AATCTACACAGAGAAATTAATCTTGCCACTATTTCATTTTGTGATTTTATTA 315
QY 1035 ttagttatagtttagagttatccaaaggttaaggtttacccaaggttttagtga 1094
DB 316 TGTATTTTAAATGTAAGAGCATCTTATTAACAAATTAATGACATATAGCTTAATTT 375
QY 1095 gggattatc---gacttagagattagtgcttactgacgacgttcaagatttttaaa 1151
DB 376 TTTATTTTTCATCTTACGAATTAATTTTGATATGAATTAATTAATTTCTGTTAAG 435
QY 1152 aatatttttttgtaacacactactattttatttttttttttttttttttttttttt 1211
DB 436 TCACATTTTAATCCACCAATTAATCTTTTATTTATTAATGACATTTTATACAAAAT 495
QY 1212 cataatataatctaactcactcgtttcactatataagttgacgttgaacatattttt 1271
DB 496 TGTCTTAATTCATTTTCTCAAAAATTAAGCACTCATATGCAATTAATATCATTAATTA 555
QY 1272 tgttaaaaaaaatgtcactttagaattcgaatcgaataattattttttcagctaa 1331
DB 556 TCTACGATATGTGATTCATTAATTAATTTGTTTATATGAATTAATAATTTCTTATTT 615
QY 1332 aatttaattgcaagtgatgactctataaataattttatttttttttttttttttttt 1391
DB 616 AATATATTTCCGATGATTTTATTAATTTTATTAATAATCTATTAATCTAATTAATGA 675
QY 1392 gtcaaacatgtgtaatttaagaaacttaatttatttttttttttttttttttttttt 1451

DB 676 GTTATTAATTACACTTTTTCATTTAGATAAATAATATCTAATTAATTTTTCGACATTC 735
QY 1452 gtaaaatgtcaagaagaaattattttagaacgaattgagtaatttttttttttttttt 1511
DB 736 GTTGTTCCTCCGAAAACATAGACCAATTAATTAATTTATTCACATTTTATTAATTA 795
QY 1512 ttaaaaa--tatgataatgaataacacaaattttattgtatgatagaacctaataatc 1569
DB 796 TGTATATATTTTTCAAAAAATTAATCAATGAAAAAATAAATTAATTAATCAAAATG 855
QY 1570 atcctaagaagtgaaacgaagataagtcacgttttggggaagctaaactatggcca 1629
DB 856 ATTTACTAAATCTGATTAATTAATTTTAATTAATTAATTAATTAATTAATTAATTA 915
QY 1630 aagtcataaaactttcttgatattatccaaatcccttaacaaatttagttaggttaata 1689
DB 916 TATTAATTAACAGATTAATGTTATTAATTAATTAATTAATTAATTAATTAATTAAT 975
QY 1690 gaccacaacatgattatcatcatatagataatcttaaaaaacttaactgacgaataat 1749
DB 976 AAGATTCGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1809
QY 1750 aaatcttcttattatcaaaatcccttaataaaaaacttattatataactaaacaat 1869
DB 1036 AAGTTTATATATTTCAAAATTTTAATTAATTAATTAATTAATTAATTAATTAATTA 1899
QY 1810 ttaatttaaaagaaataagggacacatgataataaaataatatttttttttttttt 1929
DB 1096 ATATTAACATTAATTTGATTAATTAATTTTAATTAATTAATTAATTAATTAATTA 1990
QY 1870 agtgaataattataataccagtcataatatttcaaaaaataataatttttcgta 1929
DB 1156 CACATTCATTAATTAATTTTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1990
QY 1930 gtccgataatcattactataaattcaataaaacacatgtagatgattattttatat 1989
DB 1216 ATATTTAGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1275
QY 1990 atatatat 1997
DB 1276 ATGTTTTT 1283

RESULT 5
US-08-107-755A-8
Sequence 8, Application US/08107755A
Patent No. 5721352
GENERAL INFORMATION:
APPLICANT: Moyer, Richard W.
APPLICANT: Hall, Richard L.
APPLICANT: Gruidl, Michael E.
TITLE OF INVENTION: No. 5721352e1 Entomopoxvirus Expression System
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESS:
ADDRESSEE: David R. Saliwanichik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: Florida
COUNTRY: U.S.A.
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/107,755A
FILING DATE: 19-AUG-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/827,658
FILING DATE: 30-JAN-1992
PRIOR APPLICATION DATA:

REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMM
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149

```

? INFORMATION FOR SEQ ID NO: 14
?
? SEQUENCE CHARACTERISTICS:
?
?   LENGTH: 7218 base pairs
?
?   TYPE: nucleic acid
?
?   STRANDEDNESS: single
?
?   TOPOLOGY: linear
?
?   IMMEDIATE SOURCE:
?
?   CLONE: pTZgpc-F15
?
? US-08-232-463-14

```

Query Match	3.0%;	Score 86:6;	DB 1;	Length 7218;
Best Local Similarity	3.7%;	Pred. NO. 3.2e-07;		
Matches	14;	Conservative 245;	Mismatches 124;	Total 369

QY 2443 ccgctcttgagatcttcacatgacccctctcttgaccggcatalgtcaagcatatgtctca 2502
 Db 1075 YY 1133
 QY 2503 ccgattgtgatactactaacctctctctctcaagctccttaagtctcctctaac 2562
 Db 1135 YY 1194
 QY 2563 aggcctcctcaatgctctctcaatgctcctcaaggctcctcttaatgctcttaa 2622
 Db 1195 YY 1254
 QY 2623 atgcctcttaacagcgcctcttaacatgctcctctcaaggcctcaaggcctctaac 2682
 Db 1255 YY 1314
 QY 2683 aggcctcttaacatgctcctcttaagcgcctctcgaggctcctcaccgctcaatgat 2742
 Db 1315 YY 1374
 QY 2743 ttgactattgacagaattatctcaagtalgtacgtcctcaggaggtttagtttttctt 2802
 Db 1375 YY 1434
 QY 2803 gttttaaattgtgttatatt 2825
 Db 1435 GTACCAATCTCTATCTCTTT 1457

RESULT 8

US-08-213-419B-3/c .
; Sequence 3, Application US/08213419B

; Patent No. 6333406
; GENERAL INFORMATION:

APPLICANT: Inselburg, J. et al.
TITLE OF INVENTION: CRYSTAL

GENE ENCODING PROTEIN ANTIGENS OF PLASMODIUM FALCIPARUM
AND USES THEREFOR

FILE REFERENCE: JI-002CNC
CURRENT APPLICATION NUMBER: US/08/213.419R

CURRENT FILING DATE: 1994-03-14
PRIOR APPLICATION NUMBER: US 07/920,502

NUMBER OF SEC ID NOS: 33

```

; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 1

```

; LENGTH: 6124

```

; TYPE: DNA
; ORGANISM: Plasmodium falciparum

```

```

; - FEATURE:
; - NAME/KEY: CDS

```

NAME/KEY: CDS

LOCATION: (2407)..(2439)

WHEEL/REI: CDS
LOCATION: (2598)..(3404)

NAME/KEY: CDS
LOCATION: (3580) / 3720

NAME/KEY: CDS
LOCATION: (3850) (5000)

08-213-419B-3

Query Match	3.0%;	Score 84.4;	DB 4;	Length 6124;	{
Best Local Similarity	43.8%;	Pred No. 7.5e-07;			
Matches 737; Conservative	0;	Mismatches 915;	Indels 29.	Canc 9	

[illegible]

OY	1229	ctatcctggttcacataagtgctcgtgtaaacattatttttggtaacaaaatgct	1288
Db	1665	AAAAA-----ATTGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTA	1613
OY	1289	cacttgaattcccaatgcaaaaatttatatttttcctgcgtacaaaatgaatgc	1348
Db	1612	TAAATTTTCTTCTTCCTCCAAATTAAGTTTTCATACAGATTATATAGTGT	1553
OY	1349	atgacttataataatttttatttattatcccaatgctatattgtgccaacagtgta	1408
Db	1552	ATTATATTCAATTAATTTTCTGAGGTTAAAAAAAATTAAATTTATATATCCAT	1494
OY	1409	aatgaacttaattatattcatttttcttatttttcttcaatcgtgtgtaaaatgc	1468
Db	1493	AATATATCGAACAATAGCTGATTTTCTTCTTGCTCGAGAAAATCATATACATA	1434
OY	1469	aaattattgaagaacgattggaataatttgtttcattttttaaagaatcgaata	1528
Db	1483	ATATATTATACAAAAGAA-----AATCATATATCTTATTTTACTAGACCTTATTTA	1380
OY	1529	tgaataacacaatttatgtatgtaagcaactaaatctccttaagaagtggaagc	1588
Db	1379	TTTTATTTTATTTATTTATTTCTTAATATTAACAATAATATTAAGGTCATGAATTAATTT	1320
OY	1589	aagaataagicaacgcttttggggaagctaacatgcccgaagcacaatccttct	1648
Db	1319	GCTTTTCCATCATTAAGTCTTACTAATTTTATGTTCAATGTAATTTTAAATAATACA	1260
OY	1649	tgattatcaaaaatccttaacaaatttgatagtgtaagtaagcaacaacatgattc	1708
Db	1259	TTTTATCTTAAATATGATTTTATGAGATACCTTTTATCTCTTAATATGAAAAT	1200
OY	1709	atcattatgaatatcttaaaaattacaggaataataaactcttcttattat	1768
Db	1199	ATATGATGAGAAAATAATTTATCTTAGTACTTACCATTATATATATTTTATATAA	1140
OY	1769	caaaatccttataaaaacttattatata-atactaaacaatttaattaaagaanaa	1827
Db	1139	GCATTTAAATACATTTTACATAGTAGGTCCTTAATAAATATGTTGGTAATACACCC	1080
OY	1828	agggaacacatgatacaataaataatgt-tattcttaagaatgataatataat	1886
Db	1079	AATCATTTTAAATTTATTTCTATATATATGTAATATATATATAGCTAAATTAAG	1020
OY	1887	ataccagtcacatattatcaaa-----ataaataatttttgcgtagtcgaataca	1941
Db	1019	ATGAAAAATACATATTAGTATATAGCATTTACATATATGACATTTTTCATTTCT	960
OY	1942	ttactataaatlcaataaacacatgtagatgataatttattatataatataataa	2001
Db	959	CTTTTAATATATTTTAAATAAATAATATATATAGTTTCATTAACGTADAAATATATATAT	900
OY	2002	a 2002	
Db	899	A 899	

RESULT 9
US-08-471-791-27/c
Sequence 27, Application US/08471791
Patent No. 5723595
GENERAL INFORMATION:
APPLICANT: Thompson, Gregory A
APPLICANT: Knaut, Vlc C
TITLE OF INVENTION: Plant Desaturases-Compositions
TITLE OF INVENTION: and Uses
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Calgene, Inc.
STREET: 1970 Fifth Street
CITY: Davis
STATE: California

```

COUNTRY: USA
ZIP: 95616
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 2.0 MB storage
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 7.1
SOFTWARE: Microsoftword 5.1 (a)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/471,791
FILING DATE: 6-JUNE-95
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/762,762
FILING DATE: 16-SEPT-1991
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/01746
FILING DATE: 14-MAR-1991
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/615,784
FILING DATE: 14-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/567,373
FILING DATE: 13-AUG-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/494,106
FILING DATE: 16-MAR-1990
ATTORNEY/AGENT INFORMATION:
NAME: Lassen, Elizabeth
REGISTRATION NUMBER: 31,845
NAME: Donna E. Scherer
REGISTRATION NUMBER: 34,719
NAME: Carl J. Schwedler
REGISTRATION NUMBER: 36,924
REFERENCE/DOCKET NUMBER: CGNE 69-5
TELECOMMUNICATION INFORMATION:
TELEPHONE: (916) 753-6313
TELEFAX: (916) 753-1510
TELEX: 350370 CGNE
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 3440 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: genomic DNA

```

	Query Match	2.9%	Score 83.2;	DB 1:	Length 3440;
	Best Local Similarity	52.2%;	Pred.	No. 1.le-06;	
	Matches 239; Conservative		0;	Matches 208;	Indels 11; Gaps 2
OY	1102 tgacttagattagtggtttctacgacgcgtccaaagtatctttaaataattttc				1161
Dd	457 TAAATTTTAAAAAGCTGTGTTTTTGCGTGGTAGATTTAATAATCCATATAAATCCT				398
OY	1162 ttgtacaacctaccattttaattaatttttaaccttttatcataaaaacatatana				1221
Dd	397 TTGTATTGTCACCAACAATTTCTAACGCATTAATTTTTTTTATTATTAAAAMGCTGA				338
OY	1222 tttaatactcacatcgcttcacatatnaagytcacigtgaacatatatttttgtccaana				1281
Dd	337 GAATTTTTCGTTTTTTTTTTTAATT-----CTGTATCACAAAAATATTGTTACACAA				285
OY	1282 aaattgcaactltagaattcaaigycanaaalattatlattltsagctaanaaltaatgc				1341
Dd	284 AAAAGTCACTTTAATAATTTCAAACAANAATTTACTATTATTCACGTGGAATTAATGT				225
OY	1342 aaaagcatlgctctataaataattttattatatctocaatytcataattggccaacatg				1401
Dd	224 AAACGACATGATTTTAAAACGACGAGACCCTGACCGGCCGCACGCGCATAGAANT				165

OY 1102 tgaactagatcttcgtglttctacgcagcgttcaagaattttttaaaaaafattttt .1167
| | | | | | | | | | | | | | | | |
Db 457 TAAATTTTAAAAAGCTCTGTGTTATTGGTGCTGTGTTTTTAATTCOCATAAATCTT 398

STRANDEDNESS: si
TOPOLOGY: linear

Query Match 2.9%; Score 82.2; DB 3; Length 16442;
Best Local Similarity 60.5%; Pred. No. 2.2e-06;
Matches 135; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 2514 tctactcaactctcctcctcaagagctccttaacatgctcctcctcaagagctcctta 2573
DB 16192 tct 16251

QY 2574 catgctcctcaagctcctcctcaagctcctccttaagctcctccttaagctcctta 2633
DB 16252 cct 16311

QY 2634 catgctcctcaagctcctcctcaagagccctcctcaagagccctcctcaagagccctta 2693
DB 16312 cct 16371

QY 2694 catgctcctcaagctcctcctcctcaagagctcctcctcaagagctcctcctcct 2736
DB 16372 cct 16414

RESULT 12

US-08-883-795A-36/C
; Sequence 36, Application US/08883795A
; Patent No. 5985607
; GENERAL INFORMATION:
; APPLICANT: Delcuve, Genevieve
; TITLE OF INVENTION: Recombinant DNA Molecules and Expression
; NUMBER OF INVENTION: Vectors for Tissue Plasminogen Activator
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BERESKIN & PARR
; STREET: 40 King Street West
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5H 3Y2
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/883,795A
; FILING DATE: 27-JUN-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Gravelle, Micheline
; REGISTRATION NUMBER: 40,261
; REFERENCE/DOCKET NUMBER: 7841-062
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 364-7311
; TELEFAX: (416) 361-1398
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 665 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; IMMEDIATE SOURCE:
; CLONE: Rh 32
; US-08-883-795A-36

Query Match 2.8%; Score 80.6; DB 2; Length 665;
Best Local Similarity 46.8%; Pred. No. 2.4e-06;
Matches 286; Conservative 0; Mismatches 324; Indels 1; Gaps 1;

QY 1215 aatataatttaactccatcctcgtttcattatgaagtgcatgttaacattattttgt 1274
DB 626 AATACAAATTCGAGATCAATTAACCGCTTTTAAATTTTCCGCTTTTAACTA 567

QY 1275 taacaaaaaattgcaactttagaattccaaatgcaaaatttatttattttcagctaaat 1334
DB 566 TTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 507

QY 1335 taatgcaaaagcatttgctccttaataaattttatttctcaaatgctatgctc 1394
DB 506 ATTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 447

QY 1395 aaacatgttaattatagaacttaattatatttatttcttcaactcgtgta 1454
DB 446 TAAATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 387

QY 1455 aaatgtcaaaagtaaaatttattttagaagaattgagtaataatttgcattttta 1514
DB 386 TAAATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 327

QY 1515 aaagatcgaatagaaatgaacacatttttgaatgagaaactaaatcctc 1574
DB 326 TAAATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 268

QY 1575 aagaagtgaaagcaagaatgaagcaagcttttgaggaagcaactaagcccaagtc 1634
DB 267 AATAAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 208

QY 1635 atcaaaatcttctgtattatcaaaatccttcaaaattgttagagtaagacaa 1694
DB 202 AATAAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 148

QY 1695 aacacatgattatcatcatatagaaatcttcaaaatctcaaggaataaataaact 1754
DB 147 AATAAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 88

QY 1755 ttctttattatcaaaatccttcaaaatcttcaaaattgttagagtaagacaa 1814
DB 87 AATAAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 28

QY 1815 ttaaaagaaa 1825
DB 27 AATAAACAACA 17

RESULT 13

US-08-323-170B-1/C
; Sequence 1, Application US/08323170B
; Patent No. 5733772
; GENERAL INFORMATION:
; APPLICANT: Williamson, Kim C.
; TITLE OF INVENTION: Cloning and Expression of Plasmidium
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/323,170B
; FILING DATE: 13-OCT-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/010,409

FILING DATE: 29-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Oline, Jonathan A.
REGISTRATION NUMBER: P-41,261
REFERENCE/DOCKET NUMBER: 015280-113100US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 9636 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
MOLECULE TYPE: linear
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 149..9556
US-08-323-170B-1

Query Match 2.8%; Score 79; DB 1; Length 9636;
Best Local Similarity 59.6%; Pred. No. 7.4e-06;
Matches 133; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY 2507 tcttgatctacttaccctctctctcaccaggtcccttcaatgctccttcaacaggc 2566
DB 1492 TTTTTCATCATATATACCTCTTCTTCACTACCTCTTCACTCTTCTTCACTAC 1433
QY 2567 tcccttaagtcctcttaccctctctcaccaggtcccttcaatgctccttcaacaggc 2626
DB 1432 CTCTTCACTACCTCTTCACTACCTCTTCACTACCTCTTCACTACCTCTTCACTAC 1373
QY 2627 tcccttaagtcctcttaccctctctcaccaggtcccttcaatgctccttcaacaggc 2686
DB 1372 CTCTTCACTACCTCTTCACTACCTCTTCACTACCTCTTCACTACCTCTTCACTAC 1313
QY 2687 ccccttaagtcctcttaccctctctcaccaggtcccttcaatgctccttcaacaggc 2729
DB 1312 CTCTTCACTACCTCTTCACTACCTCTTCACTACCTCTTCACTACCTCTTCACTAC 1270

RESULT 14
US-08-954-441-1/c
Sequence 1, Application US/08954441
Patent No. 6316000
GENERAL INFORMATION:

APPLICANT: Williamson, Kim C.
TITLE OF INVENTION: Cloning and Expression of Plasmodium
TITLE OF INVENTION: falciparum Transmission-Blocking Target Antigen, Pf230
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/954,441
FILING DATE: 20-OCT-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/323,170
FILING DATE: 13-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/010,409

FILING DATE: 29-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Einhorn, Gregory P.
REGISTRATION NUMBER: 38,440
REFERENCE/DOCKET NUMBER: 015280-113100US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 9636 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
MOLECULE TYPE: linear
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 149..9556
US-08-954-441-1

Query Match 2.8%; Score 79; DB 4; Length 9636;
Best Local Similarity 59.6%; Pred. No. 7.4e-06;
Matches 133; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY 2507 tcttgatctacttaccctctctcaccaggtcccttcaatgctccttcaacaggc 2566
DB 1492 TTTTTCATCATATATACCTCTTCTTCACTACCTCTTCACTCTTCTTCACTAC 1433
QY 2567 tcccttaagtcctcttaccctctctcaccaggtcccttcaatgctccttcaacaggc 2626
DB 1432 CTCTTCACTACCTCTTCACTACCTCTTCACTACCTCTTCACTACCTCTTCACTAC 1373
QY 2627 tcccttaagtcctcttaccctctctcaccaggtcccttcaatgctccttcaacaggc 2686
DB 1372 CTCTTCACTACCTCTTCACTACCTCTTCACTACCTCTTCACTACCTCTTCACTAC 1313
QY 2687 ccccttaagtcctcttaccctctctcaccaggtcccttcaatgctccttcaacaggc 2729
DB 1312 CTCTTCACTACCTCTTCACTACCTCTTCACTACCTCTTCACTACCTCTTCACTAC 1270

RESULT 15
US-08-998-416-186/c
Sequence 186, Application US/08998416
Patent No. 6239264
GENERAL INFORMATION:

APPLICANT: Philippsen, Peter
APPLICANT: Pohlmann, Rainer
APPLICANT: Steiner, Sabine
APPLICANT: Mohr, Christine
APPLICANT: Wendland, Jurgen
APPLICANT: Knechtle, Philipp
APPLICANT: Redischung, Corinne
TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSYPPII
TITLE OF INVENTION: AND USES THEREOF
NUMBER OF SEQUENCES: 1152
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6239264artis Corporation
STREET: 3054 Cornwallis Road
CITY: Research Triangle Park
STATE: No. 6239264th Carolina
COUNTRY: USA
ZIP: 27709

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/998,416
FILING DATE: 24-DEC-1997
CLASSIFICATION: 435

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: CH 0016/97
; FILING DATE: 31-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ. ID NO: 186:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 615 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: PAG1074RP
; US-08-998-416-186

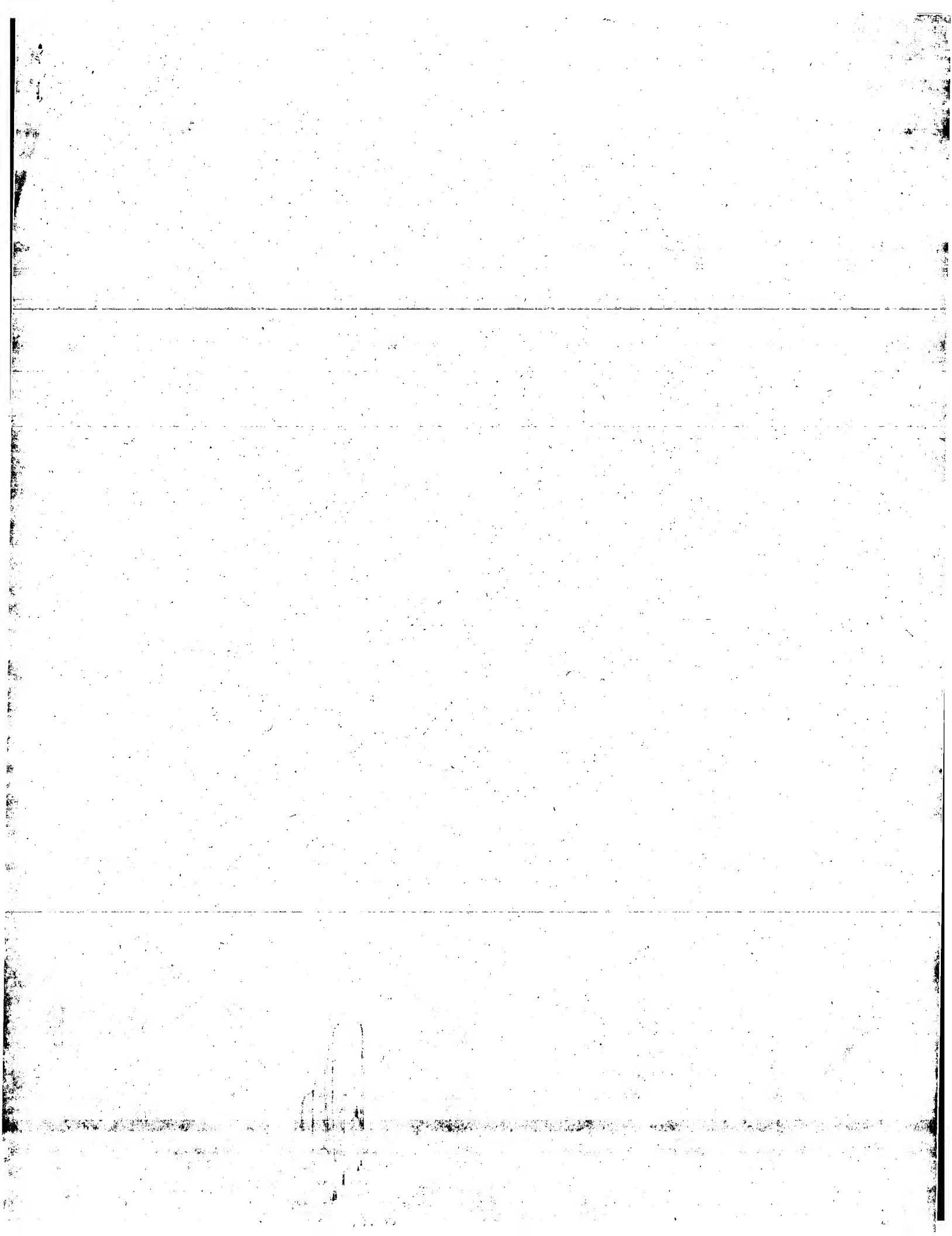
```

```

Query Match      2.7%; Score 77.6; DB 4; Length 615;
Best Local Similarity 49.5%; Pred. No. 8.2e-06;
Matches 291; Conservative 0; Mismatches 284; Indels 13; Gaps 3;

OY 1467 taaatttatttagaaacgaattgagtaataatttgcatttttaagaatcgaa 1526
DB 597 tatttaattatttaataatgataaataatttaatttaatttaatttaataa 538
OY 1527 tatgaataacacaaatttattgatagaacaaataacccatacctaagaagtgac 1586
DB 537 aatgaataaataaataatttatttaataaataaataaataaataaataaataa 478
OY 1587 gcaagaataagtaacgtttgggaaagctaataagcccaagtcacaaatctt 1646
DB 477 ttttaatttaatttaatttaattcttttaataaagtttaatttaatttaata 418
OY 1647 ctgtatttacaacatccctacaaatttagtagtaagtaagacccaacacatgata 1706
DB 417 ttttaataaataatgataatttaataaataaataatttacaataatttaataaataa 358
OY 1707 tc-----atcataatgaatctctaaataactacagcgaataatlaaactt 1757
DB 357 ttttaataaataaataatttatttttaataaataaataaataaataaataaataa 298
OY 1758 ctttatttacaacatccctataaataacttattatataactaaacaattlaacta 1817
DB 297 ataatctatttaatttttaataaataaataaataaataaataaataaataaataa 238
OY 1818 aagaataaataaggaacatgatacacaataaataatgtattcttaagaatgataa 1877
DB 237 attttaataa--ttgaacatagactaaataatgataatcaatttaataaattttta 180
OY 1878 tattaataataaccagtcacatataattatcaaaataaataataatttgcagtcgata 1937
DB 179 tatgaataatttaataa--tgatgaattgaataaatttaataaataaataaataa 122
OY 1938 atcattactataaataacataaaccacatgataatatttattatataatataat 1997
DB 121 aatgaataaataaataatttttaataaataaataaataaataaataaataaataa 62
OY 1998 ataaccctaacgccttacacacgcgaataacacataaacttctct 2045
DB 61 aatgaataaataaataaataaataaataaataaataaataaataaataaataa 14

```



Query Match	99.7%	Score: 2844	DB 6	Length 2853
Best Local Similarity	100.0%	Pred. No. 0		
Matches 2853	Conservative 0	Mismatches 0	Indels 0	Gaps 0

OY	1	ggatcccccaagaagaaacgcgaagaacaagtgttccctacccttcataataatatatttgt	60
Db	1	GGATCCCAACAAGAAAAACCAAGAACAAATGTTCTTCCTCATTAATAATATATTGT	60
OY	61	ttagcctcatcaatgtaacaaacacctttgtctcaatgafatataatgttggttttag	120
Db	61	TTGAGCCTCATCAATGTACAAAACAAATCCTTAACTTAATAGATATAAGTGTTTTAG	120
OY	121	attccaataaccggggttcgagtcatalagacttgcacattttccaacttttaaagtgtg	180
Db	121	ATTTCAAATTAACCGGGGTTCCAGTCATAGACTTGACACTTTTTCACACTTTTAAAGTGC	180
OY	181	aacgcacatctcgctacgctgtgcagtcbsggsgtgtgtgaactgcctcataatgtag	240
Db	181	AACGCACATCTCGCTACGCTGTGCAGTCBSGSGBATATGCACATGCTCATTAATATGTAC	240
OY	241	atttaaaagtgsaacccacgatacgatcgtscgctgtcgcatacgaagtgtgatgaacgtcat	300
Db	241	ATTTAAAGAAGSGAACCACAGTATCGCTGACGTGTGCATCAGGAAGATGATGCACATGCCAT	300
OY	301	attataacgtagatttgaagcttatcccttttaaatcttaataataataacaggctttt	360
Db	301	ATTATACGTAAGATTGACGTTATTCCTTTTAAATCTTAATAATAAATCACAGNGCTTTT	360
OY	361	actattaactttngcatnngtalcatacgttatacgtctcttttttgancogtltga	420
Db	361	ACTTATTAATTTTGNCAATNGTATCATAGTGTTTANGCMCTTTTTTTTGANCCGTTGA	420
OY	421	ttggtttaagcttatttgatgtatgncmacgtgaagaatgtagaacaattatattgga	480
Db	421	TTGCTTATGCTTATTITTAATGTATGNCMACGTGAAGAATGTAGAACAATTAATTATTGGA	480
OY	481	gaaaatataaatttaatatgtltaacatataabagaanaaatattatccttgaatgttacgtta	540
Db	481	GAAAAATATATTAATTAATGTTCATAATATPAGSAAANATTAATNCCTGATGTTACTGTA	540
OY	541	tggatgcgagtgtaagatcctttgaataatatttgaagacttgccttttccaazaagtaa	600
Db	541	TGGATGCGAGTGAAGATCCTTGAATAATTAATTAGAGAACTGCTTTTCCAAAAAGTAA	600
OY	601	aatatltgatalgtlaacttaagttaaacatcrgaaatlaaaaaaaatlaaatcaaaata	660
Db	601	AATATTGTATGTACACTTAAGTTAAACACATGAATAATTAANAATAATTAATCAAAAAATA	660
OY	661	gaaaaaactgtagtgcataccctcaaagctttigaactattcltgtgtccccccta	720
Db	661	GA AAAA AACTGTAGTAGATCTACCTTACAGCTTTTGAACCTTAATCTTGTTGTCACCCCCTA	720
OY	721	aacccttaagtlccaccaacacataaaattcatattgatgcgatlctataacttttagaa	780
Db	721	AACCTTAAGTTCACCAACACATAAAATTTTCATTATGTATGTCTATATCTTTTAGAAA	780
OY	781	gtgaacaaaataltatcaagitalatattatgttttccaataaagaatlaaaaaataat	840
Db	781	GTGAACAAAATATTATTCAGATTATTAATAGTTTTCAAATAAANAAGATPAANAATAATAT	840
OY	841	aaaaataatagtgttacaaaaaanaaaataatatttttccaagcgtcaanaaacac	900
Db	841	AAAAATATATGTGTGTRCAAAAAAATAATTAATTAATTAACAGCCTCANNAACAC	900
OY	901	taaaacctaacaaccttaaatllaactttagttaaaccctaacccttggataactct	960
Db	901	TAAACCTTAACCCCTAATATTTAAAGCTTTAGSTTAACCTTAACCTTTGGATAATATCTT	960
OY	961	aaacattaaacattaaaaacttaaaccttaacotcetaaaccttaaacctttagtgttta	1020
Db	961	AAACATTTAACATTAAAAACACTTAACCCCTTAACCTTAACCTTAACCTTTAAGTGTTTA	1020
OY	1021	aatgtttagtgtttttgattatagtttgattatccaagaagtttaaggtttacccaa	1080

Db	1021	AATGTTAGAGTCTTTTGATTTATAGTTTAGATTTATCCAAAGGTTTAAAGTTTATGCCAA	106
Qy	1081	gaatttctggttaaggatatagacttaaggatttagtcttaactgacgacgttcaaat	114
Db	1081	GAGTTTAGGTTTAGGATTAATACATTAGGATTTAGTGTGTTTACGACGAGCTTCAAGT	114
Qy	1141	attctttaaaaaatatttttttggtaaacacactattttattttattttttaccctt	120
Db	1141	ATTTTTTAAAAAATTTTTTTTTTGTAACAACTACATTTTTTTATTTATTTTATTTACCTTTT	120
Qy	1201	tataataaacaataataatttaatactccatccgtttcataatgaatgtaagtgcattga	126
Db	1201	TATATTAAAAACATATATATATTTAATCTCATCGTTCGTTTCATATATAGTGCATTTGTA	126
Qy	1261	cattattttttgttcaaaaaaatgttaccttttgaaatcccaatgcacaaatttattat	132
Db	1261	CATTATTTTTTTGTTCAAAAAAATGTGCATTTAGAAATTCCAATGCAAAAATTTATTTAT	132
Qy	1321	ttttcagctaaaaattaatgttccaagtgcatgcatctctataaataatttattatccca	138
Db	1321	TTTTCAAGCTAAATTTAATGTCCAAAGGCGCTTATGCTTTAATAATTTTATTATTCCTCA	138
Qy	1381	atgcctattgttgcacacgtggttaatttaataagaacttaatatatttcaattttt	144
Db	1381	ATGCTATATTGGTCACAAACATGTGTATTTATATGAAACCTTAATTTATTTATTTATTTT	144
Qy	1441	tcttaactctgtybaaaaaatgtaagaatlaaaattatttagaagaagatgtagtaattt	150
Db	1441	TCTTATCTGTGTAAAAAATGTCAAGGTAAATTTATTAGAAAGCATGTAGTATATTTT	150
Qy	1501	tgttcaattttttaagaatatacgataatgaataaacaacatttactgtatagtaac	156
Db	1501	TGTTTCAATTTTTTAAAAAGATATCGAATATGAAATTAACAAATTTATTGTATGATGAACC	156
Qy	1561	taaaaaatlcactcctaagaagtgtaacgcgaagaataaagtcacgttttggggaaagctaac	162
Db	1561	TAAAAATTCATCTCTAGAAGGTGAACGCAAAATTAATGTCACAGCTTTGGGGAAAGCTAAC	162
Qy	1621	tatgcccccaagtcatacaaatcttctctgtattattcaaaatcccttaacaatttagtta	168
Db	1621	TATGGCCCAAACTCAACAAATCTTCTCTGTATTTATCAAAATCCCTTACAAATTTAGTTA	168
Qy	1681	gagttaatagaccaacaacatgatatacatataggaatatttcaaaaaattactagc	174
Db	1681	GAGTTAATAGCAACAAACATGATTAATCATATTTTGAATTTCTTAAAAAATTTACTAAC	174
Qy	1741	gaataatnaaaatcttcttatttatttcaaaatcccttaaaaaactttattatataac	180
Db	1741	GAAATATTTAAAAATCTTCTTTATTTATTCAAAAATCCCTTATTTATTTATTTATTC	180
Qy	1801	taaaacaattttaatttaaaaaaataagaagccatggacatgagatacaataataatgta	186
Db	1801	TAAACCAATTTTAAATTTAAAGAAAAATATAGGACCATGAGTACATAAAAAATATATTTGTTT	186
Qy	1861	tccttaagaatgataataattataatactacagctcatatattatcaaaaataataata	192
Db	1861	TCTTAAAGTACTGATATATTTAATATATATTCACATCATATTTTATCAAAATTAATTAATA	192
Qy	1921	tttttcgtagccgataataactataataatcatataaacaacatgtagatgtaattt	198
Db	1921	TTTTTCGTAGCCGATATATATTAATTAATAATTCATTAATAACCAACATGATGATATATTT	198
Qy	1981	tattataataataataataaacccttaagcccttaccactgataaccatacaaatctt	204
Db	1981	TATTTATATATATATATATTAACCCCTTAACGCCCTTACCATCTCATATACATCAAAACCTTT	204
Qy	2041	cttcctgcttcgccttaacctcaagctctcgaaagttaaaaaaacaatgaagaatgtcaaac	210
Db	2041	CTTCTGCTTTGCTTAACACTCAAGGCTTCGAAAGTAAAAAACAATGAAGATGTCAAC	210
Qy	2101	tgttctctgctatgataccctcttcttaagctctgttcaacatccaagaattacagcaacagac	216

DB	2101	TTGTTCTTGATGATGCTCTTTAAAGCTGAGTCATCCAAAGTTACAGAACAGANAC	2160
Qy	2161	tagagtcataactaactaacaagaagctctctccatctcgcggaacttactctgtcttcaeccca	2220
Db	2161	TAGAGTCATCAACAAACCAAGAGCTCTTCCATATCCGCGGACATTACCTCGCTTTCACCCCA	2220
Qy	2221	agcaaatctggccgtctccgtgtgtctccggaaaaagccctccctcgagaggaacttcgactaa	2280
Db	2221	AGCAACATTGGCCGCTCCGCTGTGGCTCCGGAAAAAGCTTCCCTGCAAGCCACTTCGCACTAA	2280
Qy	2281	ctccgcttccatctgcacaaagaagtlcacacagatgctltgaagacaagaagaagtagta	2340
Db	2281	CTCCGCTTCATCTCGACACAGGAAGTCAACAGATGCTTTAAGACAGAAAGAGAGTAGTA	2340
Qy	2341	catgtttaaataatctgcgtgagactctcttcacacagaagaagccgctatlgatcgaaat	2400
Db	2341	CATGTTTAAATGAATATCGTATGACTTTCTTCACACAGAAAGCCGCTATTTGATCGGAAT	2400
Qy	2401	gttcgacgcgataaagaagatlgacaagaatttggagaagaacgctctttgatcttcc	2460
Db	2401	GTTGGCGCGCATCAAGAAATGAACAAAGATGTGAAGAAGACCGCTTTTGGATCTTTCC	2460
Qy	2461	atgacccctcttgacccgagcatgtcaagctacatctgttccacogtltgtatctact	2520
Db	2461	ATGACCCCTTCTTGACCGGCTATGTCAAGTCAATGCTCCACCGTGTGTGATCTACTT	2520
Qy	2521	caactccctcttcaacagagctccttataatgtctcctcttcaacagagctcctacatgtctc	2580
Db	2521	CACCTCCCTCTTTCACAGAGCTCTTTACATGCTCTTTCACAGAGCTCTTTCACATGCTC	2580
Qy	2581	cttcaatagctccttcaagaagctcctttaaattgtccctttaaagtctcttcaatgtctc	2640
Db	2581	CTTCAATAGCTCTTTCACAGAGCTCTTTAAATGCTCTTTAAATGCTCTTTCACATGCTC	2640
Qy	2641	cttcaatagctccttcaagaagcccttcaacagcccttcaacagcccttcaatgtctc	2700
Db	2641	CTTCAATAGCTCTTTCACAGAGCCCTTTCACAGAGCCCTTTCACAGAGCCCTTTCATGCTC	2700
Qy	2701	cttactgcgccctctgcagagctccttcaacagcgctcaatgtattagctattatagaat	2760
Db	2701	CTTACTGCCCTCTTGCAGAGCTCTTTCACAGGCTATTTAGCTATTTGATGATAGAT	2760
Qy	2761	tattcaagatattgatctcttgaggagtttagtttttcttctgtttaaatttggtt	2820
Db	2761	TATTCAAGATTTGATGATGCTCTTGCAGAGGTTTATGTTTTTTCTTGTAAATTTTGTGTT	2820
Qy	2821	tattttgagaagaacgctcttgcattttaact	2853
Db	2821	TATTTTGAGAAAACCGTCTTTGATTTTAACCTT	2853
RESULT	2		
AF136223		2966 bp	DNA
LOCUS			linear
DEFINITION			PLN 23-SEP-1999
ACCESSION	AF136223		
VERSION	AF136223.1		
KEYWORDS			
SOURCE			
ORGANISM			
	rape.		
	Brassica napus		
	Euarysta; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;		
	Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.		
	1 (bases 1 to 2966)		
	Fourgoux-Nicol, A., Drouaud, J., Haouazine, N., Pelletier, G. and		
	Guerche, P.		
	Isolation of rapeseed genes expressed early and specifically during		
	development of the male gametophyte.		
	Plant Mol. Biol. 40 (5), 857-872 (1999)		
	JOURNAL		
	MEDLINE		
	PUBMED		
	10487220		
	2 (bases 1 to 2966)		
	Guerche, P., Fourgoux-Nicol, A., Drouaud, J., Haouazine, N. and		
	AUTHORS		

TITLE	Direct Submission
JOURNAL	Submitted (19-MAR-1999) Station de Genetique et Amelioration des Plantes, Institut National de la Recherche Agronomique (INRA), Route de St-Cyr, Versailles 78026, France
FEATURES	Location/Qualifiers
SOURCE	1..2966 /organism="Brassica napus" /db_xref="taxon:3708"
mRNA	2022..2966 /product="M3.4.protein"
CDS	2085..2741 /function="unknown" /note="putative cell-wall protein; expressed specifically in microspores" /codon_start=1 /product="M3.4.protein" /protein_id="AAD24197.1" /db_xref="GI:4574746"
	/translation="MKNTVLYLAILPEISCTSKVTATELESSTNOELFSLRHPRIH PKOHWPRGSGKAPAGIFRLTPPHLPDEIVRLCLNDKRKEVGTENDLAIEFTFKALD GSECCAIKKNNKDCEKTVFGSFDPDLPTGLKHSVVGSTSPPSQAPIAIPSSQ APSHAPSAPSQAPLNPIANPILNAPLHAIPSAQSAPSQAAPLHAPILPPSQAPSPAQ"
BASE COUNT	1002 a 518 c 389 g 1057 t
ORIGIN	
Query Match	99.6%; Score 2840.8; DB 8; Length 2966;
Best Local Similarity	.99.6%; Pred. No. 0;
Matches 2842; Conservative	0; Mismatches 11; Indels 0; Gaps 0.
OY	1 ggaaccacaagaacccaagaaagcaaatgttcttaacctcataatatattg
Dd	1 GGATCCCCAACGAACCGAGAAGCAAAATTGTTCTTCACCTTCAATAATATATTGGT
OY	61 ttcaagccatcaaatgataacaacactccttgaccctaagtataaigtgtgttag
Dd	61 TTCACCTCATCAATAGTATACAAAACATCCTTAGCTCAATGATATAAAGTGTGTTAG
OY	121 attcaataaccgggtttcgagtcatagaactgcaactttcacacttttaaaagt
Dd	121 ATTCAATRACCAGGGGTTCGACTCATAGACTTGACACTTTTCACACTTTTAAAAAGTGG
OY	181 aacgcacatlcgtcgaegtgcgatcagaagtgatgcaactgctctataatagt
Dd	181 AACGCACATATCGTIGACGTGTCGATCAGAGTGAAGTGAACACTGCTCATTAATATGTTG
OY	241 atttaaagtgaaaccaacagatcgtcgaegtgcgatcagaagtgatgcaactgcat
Dd	241 ATTTAAAGTGAACCCACAGTATCGCTGACGTGCAATCAGAGTGAATGACAATGCCAT
OY	301 attataagtagattgagcttatcccttttaaaccttaataataacagngcttt
Dd	301 ATTATACGTACATTTGACGTATTTCCTTTTAAATCTTAATATATATATATATATATAT
OY	361 acttaataatttggcatngtatacatagtttcatlfgcnctcttttttgtanccgtlga
Dd	361 ACTTATTAATTTTGGCATNGTATTATCATNGTATTATGCTCTCTTTTTTTTGAACGGTTGA
OY	421 ttggtttatgcttatttgaagtlnucncaaglaagaaatlgaaacaattatattgga
Dd	421 TTGGTTTATGCTTATTTGCAATGTACCAACGTATACGAATATGCAATTTTATTTTGA
OY	481 gaaaataataattataatgttcaataataagaagaaataatttnccttgatgttactga
Dd	481 GAAAATATATATTTAATATATGTTCAATATATATAGGAAATATATATATCTTGATATATCTGA
OY	541 tggatgagagtagaagatcttgaataataatttgaagaacttgccttttccaaaaagta
Dd	541 TGGATGCGAGTAGAAGATCTTTGCAATATATTTGACAACCTTTCCTTTTCTCAAAAAGTAA
OY	601 aatatttgatgtacttaagttaacacatgaaatttaaaaaaaaaaatttaatacaata

D	601	AATATTGATGTGTAACCTTAAGTTAAACATGAAATTTAAAAAAATTAATTAACAATA	660
O	661	gaaanaactgttagtgaatcacccttcaacglttgaacttaattcgttccaccctta	720
D	661	CAAAAATCGATGATGTGATGATACCCCTCAACGTTTGAAGACTTTCCTGGTTCACCCCTA	720
O	721	aacctctagctcccaacaacaatcttcattcttcatactatcctttgaaa	780
D	721	AACCTTAAGTTCACCAACAAATAAATTTCAATTATTCATATCTCATCTTTTAAAAA	780
O	781	gtgagacaaatatatcaagtatactagtlttccaataaagaataaataaat	840
D	781	GTGAACCAAAATATATACCAAGTTATATATGTTTTCAAAATAAAAAGATAAAAATTAAT	840
O	841	aaaaataatagtaagtatacaaaaaaaaataatattttaccagcgcaaaaaaac	900
D	841	AAAAAATATATAGTATACAAAAAAAATAATTAATATTTTACCAGCGCAAAAAAC	900
O	901	taaaacctaaaccttaataacttttagttaaacccttaacccctttagataactt	960
D	901	TAAAACCTAAACCCCTAAATATTAACCTTTAGTAAACCCTAACCTTTGGATTAATCTT	960
O	961	aaacatnaaacatnaaacacaaaccccaaacctcaactctaacccttaagtgtta	1020
D	961	AAACATTAACCATTTAAACACTAAACCCCTAAATCCCTAACCTTAACCCCTTAAGTGTTA	1020
O	1021	aatgttagtgttttgatllatagtttaggattctcaagaagtttaagtttaccoc	1080
D	1021	AATGTTAGGTGTTTTGATTTATATGATTAGATTTATCCAAAGGTTTAAGGTTTACC	1080
O	1081	gagttatggttaggattagacttaggattagtggtttacagacggttcaaggt	1140
D	1081	GAGTTATGCTTTAGGAGTATGACTTATGAGTTATGAGTTTACTGACGACGCTTAACGT	1140
O	1141	attttttaaanaaatatttttttgaacaactactatttttatttttttttaccctt	1200
D	1141	ATTTTAAAAAAATATTTTTTTTGAACAACCTACTATTTTATTTATTTTTTACCCTTT	1200
O	1201	tatatnaaacatatataatataaataaccccatcctgtttcatatgaatgaatgtca	1260
D	1201	TATATTAANAACATATATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	1260
O	1261	catatatttttgttacaanaaaatgtccactttagaatcaatgaagcaaaatttat	1320
D	1261	CATATTTTTTTGTACAAAAAATTCCTACTTAAATTTCCAAAGCAAAATTAATTAAT	1320
O	1321	tttttaagctaaataatattgcaagtgcatctgactcttaanaataattlatatc	1380
D	1321	TTTTTAAGCTAAATTAATTTGCAAGTGCAATGATCTTAAATTAATTTATTTATTC	1380
O	1381	atgcataatgttgaacaacatgttbaattatagaacacttaattatcttatttt	1440
D	1381	ATGCATATTTGCTCAACATCTGTATTAATTAATTAATTAATTAATTAATTAATTTT	1440
O	1441	tccttaactgtgtaaaatgttcaagtaaaattatattggaagaacttggtatatt	1500
D	1441	TCTTAATCTGTGTAAAAATGTCAAAAGTAAATTTATTAGAAACCAATGTGATTAATTT	1500
O	1501	tgtttacattttttaagaatatacgaatatgaaataacaacttattatgtatgaac	1560
D	1501	TGTTTCATTTTTTAAAGATATCGAATATGAATTAACCAATTTTATTTGATGATGAAC	1560
O	1561	taaaatctatcctaaagaagtggaagcagaataagctcaacglttggggaagctaac	1620
D	1561	TAAAAATTCATCTTAAGAAAGTGAACGCAAGATATAGTCACACGTTTGGGGAAGCTAAC	1620
O	1621	tatggccaagatcatcaaaaactttctgtattatccaanaatccttaacaattagtta	1680
D	1621	TATGGCCCAAGATCATCAAAATCTTTCTTGATTAATTAATTAATTAATTAATTAATTAAT	1680
O	1681	gagttatagaccaaacacatgattatatacattatgaatatcttaaaaatttactag	1740

[illegible]

2821 tatttgagaaacgctcttgatlttaactt 2853
|||||
Db 2821 TATTTGAGAAACGCTCTTGATTTTAATT 2853

RESULT 3
A97605 A97605 674 bp DNA linear PAT 26-JAN-2000
LOCUS Sequence 2 from Patent WO9915678.
ACCESSION A97605
VERSION A97605.1 GI:6780906
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 674)
AUTHORS Pelletier, G. and Drouaud, J.
TITLE MICROSPORE-SPECIFIC PROMOTER AND METHOD FOR OBTAINING HYBRID PLANTS
JOURNAL Patent: WO 9915678-A 2 01-APR-1999;
AGRONOMIQUE INST NAT RECH (FR); PELLETIER GEORGES (FR)
FEATURES
source 1..674
location/Qualifiers
BASE COUNT 155 a 181 c 124 g 214 t
ORIGIN

Query Match 16.2%; Score 462; DB 6; Length 674;
Best Local Similarity 81.9%; Pred. No. 3e-53;
Matches 614; Conservative 0; Mismatches 35; Indels 101; Gaps 3;

2105 tcttgatgatctcttcttcttgaagctggtgacatcaaaagttaacgcaacagactaga 2164
|||||
Db 1 TCTTGATGATTTCTTATGATGATGTGCATCATCAAGTACAGCAACAGACTAGA 60

2165 gtcataaataacaaagagctctctcatcgagcaacttaacctgttcaacccaaga 2224
|||||
Db 61 GTCATAACTAACCAAGAGCTCTTCTATCGCGGCACTTGCTGCTTTCACCCCAAGCC 120

2225 acattgagcgttccgtggtcgcggaagacgtccctgcagcgaacttccgaactaacc 2284
|||||
Db 121 ACATTGCGCGTCTGTGCGTCCGGAAGGCTTCCCTGCGAGCGCACTTCGACCAACTGC 180

2285 gtccatctgcgcaagaaagtaacagatgttgaaagcaagaagaaaggaagtagatcaatg 2344
|||||
Db 181 GTTCCATCTGCGCACAGGAAGTCAACAGATGCTTGTCCGACAAAGAGAGGTAGTACATG 240

2345 tttaatgatatcgctgagaaacttcttcccaagaaagccgctattggaatcggaatgttg 2404
|||||
Db 241 TTTTGTGATGATGCTTGAGACTTTCTTCACACAGGAAGCCGTTATTTGATCGGAATGTTG 300

2405 cgcgcgatacaagaagatgaacaagaatgttgaaagacgcttcttgatcttccatga 2464
|||||
Db 301 CGCGGAGATCAAGAGATGAACAAAGATTGTGAGAGACGCTTTGATCTTTCATGA 360

2465 cccctcttgagccggtatgttaagtaacatgtgtccacggttgttgatcttacc 2524
|||||
Db 361 CCCCTTCTTGACAGGCTATGTCAAACTACATTCTCCACCGTGTGAGATCACTTCAC 420

2525 tccctctcaagagctccttaccatgtcccttctcaagagctccttaccatgtcccttc 2584
|||||
Db 421 TCTCTCTTC----- 429

2585 acatgctcctcaagagctcctttaaagtctcctttaaagtctccttaccatgtcccttc 2644
|||||
Db 430 -----ACATGCTCTCTTC 441

2645 acatgctcctcaagagcccttcaagagcccttcaagagcccttaccatgtctccttc 2704
|||||
Db 442 ACAGGCTCTTTACATGATCTCTTCACAGGCTCTCTTACATGACCCCTTCACAGGCTCTCTTT 501

2705 actgcccccttcgagagctccttcaacgagctcaagta-ttaagctattgatagaattat 2763
|||||
Db 502 ACTGCCCCCTTACAGAGCTCTCCACCGGCTCAGTATTTAGCTATTTGTGAATTA 561

2764 tcaagatgtatgtccatgagaggttttagtttttcttgctttttaaatttggttat 2823
|||||
Db 562 TCAAGTGTGATGCTCCTAGGAGGAGTTTAG-GTTTTCTGTTTAAATTTGTGTTTAT 620

2824 tttagaagaacgctcttgatlttaactt 2853
|||||
Db 621 TTGAGAAACCGCTTTGATCTTAATT 650

RESULT 4
A97604 A97604 497 bp DNA linear PAT 26-JAN-2000
LOCUS Sequence 1 from Patent WO9915678.
ACCESSION A97604
VERSION A97604.1 GI:6780905
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 497)
AUTHORS Pelletier, G. and Drouaud, J.
TITLE MICROSPORE-SPECIFIC PROMOTER AND METHOD FOR OBTAINING HYBRID PLANTS
JOURNAL Patent: WO 9915678-A 1 01-APR-1999;
AGRONOMIQUE INST NAT RECH (FR); PELLETIER GEORGES (FR)
FEATURES
source 1..497
location/Qualifiers
BASE COUNT 107 a 130 c 70 g 190 t
ORIGIN

Query Match 13.8%; Score 393.2; DB 6; Length 497;
Best Local Similarity 98.0%; Pred. No. 5.8e-44;
Matches 398; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

2448 tttagatcttccagaccccttcttgaccggtatgtcaagctaatgtatccacggt 2507
|||||
Db 1 TTGATCTTTCCATGACCCCTTCTTGACCGGCTATGTCAAGTACATGCTCCACCGTT 60

2508 gtgagatctactaacctcctccttcaagagctccctttaaagtctccttcaagagct 2567
|||||
Db 61 GTTGATCTACTTACCTCTCTTCAACAGGCTCTTACATGCTCTCTTCAACAGGCT 120

2568 ccttcaatgctccttcaatgtctccttcaagagctcctttaaagtctcctttaaagtct 2627
|||||
Db 121 CCTTCAATGCTCTCTTCAATGCTCTTCAACAGGCTCTTAAATGCTCTTTAAATGCT 180

2628 ccttcaatgctccttcaatgtctccttcaagagcccttcaagagcccttcaagagcc 2687
|||||
Db 181 CCTTCAATGCTCTCTTCAATGCTCTTCAACAGGCTCTTCAACAGGCTCTTCAACAGGCT 240

2688 ccttcaatgctccttcaatgtctccttcaagagctccttcaagagctccttcaagagct 2747
|||||
Db 241 CCTTCAATGCTCTCTTCAATGCTCTTCAACAGGCTCTTCAACAGGCTCTTCAACAGGCT 300

2748 tattgataaataatcaagatagatgtcctcaagagagtttagtttttcttgctttt 2807
|||||
Db 301 TATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360

2808 aaaaattgtgttattttgagaagaacgctccttgaatttaactt 2853
|||||
Db 361 AAAATTGTTGTTTATTTTGAAGAAACCGCTTTGATTTTAATT 406

RESULT 5
AB005238 75188 bp DNA linear PLN 27-DEC-2000
LOCUS AB005238/c
DEFINITION Arabidopsis thaliana genomic DNA, chromosome 5, p1 clone:MKP11.

ACCESSION	AB005238 BA000015	CDS	join(8327..8440,8574..8786,8879..9045,9427..9508,9616..9684,9797..9967) /note="gene_id:MKP11.2 unknown protein"
VERSION	AB005238.1 GI:22643310		
KEYWORDS			
SOURCE	Arabidopsis thaliana (strain:Columbia) DNA, clone_11b-Mitsui P1 clone:MKP11.		
ORGANISM	Arabidopsis thaliana		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis. 1 (sites)		
AUTHORS	Sato,S., Kotani,H., Nakamura,Y., Kaneko,T., Asamizu,E., Fukami,M., Miyajima,N. and Tabata,S.		
TITLE	Structural analysis of Arabidopsis thaliana chromosome 5. I. Sequence features of the 1.6 Mb regions covered by twenty physically assigned P1 clones	CDS	join(8327..8440,8574..8786,8879..9045,9427..9508,9616..9684,9797..9967) /note="gene_id:MKP11.2 unknown protein"
JOURNAL	DNA Res. 4 (3), 215-230 (1997)		
MEDLINE	97471969		
REFERENCE	2 (bases 1 to 75188)		
AUTHORS	Nakamura,Y.		
TITLE	Direct Submission		
JOURNAL	Submitted (02-JUL-1997) Yasukazu Nakamura, Kazusa DNA Research Institute, Department of Plant Gene Research, 1532-3, Yana, Kisarazu, Chiba 292-0812, Japan (E-mail: ynakamura@kazusa.or.jp, Tel:81-438-52-33935, Fax:81-438-52-3334)		
COMMENT	Address for correspondence: kaos@kazusa.or.jp For the latest information on annotation of this clone, please see http://www.kazusa.or.jp/kaos/cgi-bin/seq_graph.cgi?c=MKP11 'product' or 'note' qualifiers. Genes that have no significant protein similarity are described as 'unknown protein' The software programs used to predict genes include: Graal (Informatics Group, Oak Ridge National Laboratory, http://compbio.ornl.gov/Graal-1.3/), GENSCAN (Chris Burge, MIT, http://CCR-081.mit.edu/GENSCAN.html), NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of Denmark, http://www.cbs.dtu.dk/services/NetGene2/) and SplicePredictor (Volker Brenkel, Stanford University, http://gremlini.zoel.iastate.edu/cgi-bin/sp.cgi). Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, Washington University School of Medicine, St. Louis, http://genome.wustl.edu/eddy/tRNAscan-SE/). This sequence may not be the entire insert of this clone. It may be shorter because we remove overlaps between neighboring submissions. The 5' clone is F2K13 and the 3' clone is T10B6.	CDS	join(8327..8440,8574..8786,8879..9045,9427..9508,9616..9684,9797..9967) /note="gene_id:MKP11.2 unknown protein"
FEATURES	Location/Qualifiers		
SOURCE	1..75188		
	/organism="Arabidopsis thaliana"		
	/strain="Columbia"		
	/db_xref="taxon:3702"		
	/chromosome="5"		
	/clone="MKP11"		
	/clone_11b="Mitsui P1"		
	complement(1..99)		
exon	/note="CDS is reported in Acc# AP002072 9b AA26111.1 gene_id:F2K13.310"	CDS	join(8327..8440,8574..8786,8879..9045,9427..9508,9616..9684,9797..9967) /note="gene_id:MKP11.2 unknown protein"
	/number=3		
	/evidence="not experimental"		
exon	/product="protein; similar to unknown protein"		
	complement(185..680)		
	/note="CDS is reported in Acc# AP002072 9b AA26111.1 gene_id:F2K13.310"		
	/number=2		
	/evidence="not experimental"		
exon	/product="protein; similar to unknown protein"		
	complement(1306..1419)		
	/note="CDS is reported in Acc# AP002072 9b AA26111.1 gene_id:F2K13.310"		
	/number=1		
	/evidence="not experimental"		
	/product="protein; similar to unknown protein"		

OY	2233	acaagaaggaagtagtaccatggttttaatatgatcgcgcgaagctcttcacccaagaag	2382
Db	74948	ATTGCAAGAAATGTAAAAACATGTTGGCGGAGATGCCTTAAGCTTTCACCCCAAC	74869
OY	2383	cgcgcattgatccgaaagtgttcggccgcgcatacaagaatgacaaaagatttgagaaga	2442
Db	74888	CAGCATTTGGAATTGGATTGTGGCGCTCCTCATCCACAAGAATGATGAAGATTGTGAATAA	74829
OY	2443	ccgcttcttgatclttccatgtaaccccctctcttgaccgcatgtatgaactgacctca	2502
Db	74828	CTGTCTTTGGAGGCCATCCACAACCCTTTCTTGAGACTGTGTGGTTAAAGCTTCATCTCTA	74769
OY	2503	ccgttgttgatctactcacctcctcccttcaaggct	2540
Db	74768	CCAAGCTAAATCACCTCGTGTGCTCCACATCCAGCT	74731
RESULT	6		
LOCUS	AF136223/c	2966 bp	DNA linear PLN 23-SEP-1999
DEFINITION	Brassica napus M3.4 protein gene, complete cds.		
ACCESSION	AF136223		
VERSION	AF136223.1	GI:4574745	
KEYWORDS	rape.		
SOURCE	Brassica napus		
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica. 1 (bases 1 to 2966) Fourgoux-Nicol,A., Drouaud,J., Haouazine,N., Pelletier,G. and Guerehe,P. Isolation of rapeseed genes expressed early and specifically during development of the male gametophyte Plant Mol. Biol. 40 (5), 857-872 (1999)		
TITLE	JOURNAL MEDLINE	99415142	
PUBMED	10487220		
REFERENCE	2 (bases 1 to 2966)		
AUTHORS	Guerehe,P., Fourgoux-Nicol,A., Drouaud,J., Haouazine,N. and Pelletier,G. Direct Submission Submitted (19-MAR-1999) Station de Genetique et Amelioration des Plantes, Institut National de la Recherche Agronomique (INRA), Route de St-Cyr, Versailles 78026, France		
FEATURES	source	location/Qualifiers	
	1..2966	/organism="Brassica napus"	
	/db_xref="taxon:3708"		
MRNA	20221..2966	/product="M3.4 protein"	
CDS	2085..2741	/function="unknown"	
	/note="putative cell-wall protein; expressed specifically in microspores"		
	/codon_start=1		
	/product="M3.4-protein"		
	/protein_id="AAD24197.1"		
	/db_xref="GI:4574745"		
	/translation="MKNTVFLPLNAILPILSCYTSKYVTATELESSNQELFSLRHPRFH PKQWMPRGSKAPRACHFTLPTHLDEVRCLNDKREVGTCENDIAETFTTKRAAI GSCEHAIKKMNKCIEVFSGSFHDPLELTGVKLDHCSTVVGSTSPPSOAPLPAASSO APSHPASHASQAPLNAPLNALPLAIPAHAPSOAPSQAAPLPAIPLPESQAPSPAO		
BASE COUNT	1002 a	518 c	389 g 1057 t
ORIGIN			
Query Match	6.2%; Score 175.8;	DB 8;	Length 2966;
Best Local Similarity	68.8%;	Pred. No. 7.1e-15;	
Matches	311;	Conservative	0; Mismatches 113; Indels 28; Gaps 4;
OY	793	attacaagtataataatggttttccaataaaagataaaaaataataaaataaatag	852

Accession	Source	Organism	Reference	Authors	Title	Journal	Features
Dn 1230	AGTATTAATATATATATGCTTTTAATATATAAAGGTAATAAATAATATAAATACTAG	1171					
Qy 853	tagttacaaaaa-----aattaattttaccagcgt-----canaaac	898					
Dn 1170	TTGTTACAAAAAATAATATTTTAAAAAAACTTTGAAACGCGTCACTAATAACACTAAT	1111					
Qy 899	actaaacctaaaccttaaatatgaactctttagtgtaaacccctaaacctttagtaaac	958					
Dn 1110	CCTAAGTCATATATCCCTAAACCATTAACCTCTGGGTAAACCTTAACCTTTGGATTAATC	1051					
Qy 959	ttaaaccttaaacatlaaacactaaacctaaaccttaaaccttaaaccttaagtggt	1018					
Dn 1050	CTAAACATTAATCAAAAACAAACACTAAACATTTAAACCTTAAGGTTTAGGTTAGCAAT	991					
Qy 1019	taaatgtttagtgctttagttagttagttagttagttagttagttagttagttagttagt	1078					
Dn 990	TAGGCTTATGCTTTTAATGTTTAATGTTTAATGTTTAATGTTTAATGTTTAATGTTTAAT	931					
Qy 1079	aagaatttaagtttaggagatagacttaggagtttaggttagttagttagttagttagt	1138					
Dn 930	AAAAGTTTATATTTTGGGTTTAGGTTTAGGTTTAGGTTTAGGTTTAGGTTTAGGTTTAGG	879					
Qy 1139	gtaatttttaaaaaaatattttttgttaaacactactattttatttatttttttaacct	1198					
Dn 878	ATAAT-----AATTTTTTTTTTTTTTCTACTATATTTTATTATTATTATTATTATTATT	825					
Qy 1199	tttaattaaacaacataataatttaact 1230						
Dn 824	TTTATTTGAAAAACATATATTAACCTTGATTAAT 793						
RESULT 7	A97606/c						
LOCUS	A97606	2853 bp	DNA	linear	PAT 26-JAN-2000		
DEFINITION	Sequence 3 from Patent WO9115678.						
ACCESSION	A97606						
VERSION	A97606.1	GI:6780907					
KEYWORDS							
SOURCE	unidentified.						
ORGANISM	unidentified						
REFERENCE	1 (bases 1 to 2853)						
AUTHORS	Pelletier G. and Drouaud J.						
TITLE	MICROFOR-SPECIFIC PROMOTER AND METHOD FOR OBTAINING HYBRID PLANTS						
JOURNAL	Patent: WO 9915678-A 3 01-APR-1999;						
FEATURES	AGRONOMICUE INST NAT RECH (FR); PELLETIER GEORGES (FR)						
source	Location/Qualifiers						
	1. 2853						
	/organism="unidentified"						
	/db_xref="taxon:32644"						
BASE COUNT	961 a 506 c 381 g 996 t 9 others						
ORIGIN							

[illegible]

QY	1019	aaatgcttgcgtttgttgccttataagtttagagattatccaaaggtttaagtttacc	1078
Db	990	TAGCGTTTACTGTTTAAATGTTTAAAGTTTAAATTTAAATCCAAAGGTTTAGGTTTACCT	931
QY	1079	aagagttacagtttagagagattagacttagagtttagctgtttactagcagcgttcaaa	1138
Db	930	AAAAGTTTAAATATTTTGGGTTTAGGTTTAG-----TGTTTTTGACGCTGCTAATAA	879
QY	1139	gtttttttaaaaaaatcttttttgcacaactactattttatttttttttaccct	1198
Db	878	ATATTT-----AAATTTTTTTTTTTTTTAACTACTATTTATTTTATTTTATCTT	825
QY	1199	tttatataaacaatacataataatcaact	1230
Db	824	TTTATTTGAAAAACATATATATTAAGTTGATAT	793
RESULT	8		
LOCUS	DM011584/c	4601 bp	DNA linear INV 23-JUL-1994
DEFINITION	Drosophila melanogaster Oregon-R mitochondrial A+T region:		
ACCESSION	U11584		
VERSION	U11584.1	GI:508826	
KEYWORDS	mitochondrial DNA; A+T region; tandem repeats.		
SOURCE	fruit fly		
ORGANISM	Mitochondrion Drosophila melanogaster		
REFERENCE	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.		
AUTHORS	Lewis,D.L., Farr,C.L., Farquhar,A.L. and Kaguni,L.S.		
TITLE	Sequence, Organization and Evolution of the A+T Region of		
JOURNAL	Drosophila melanogaster Mitochondrial DNA		
MEDLINE	Mol. Biol. Evol. 11, 523-538 (1994)		
REFERENCE	94283822		
AUTHORS	2 (bases 1 to 4601)		
JOURNAL	Kaguni,L.S.		
TITLE	Direct Submission		
JOURNAL	Submitted (28-JUN-1994) Laurie S. Kaguni Ph.D, Dept. of		
FEATURES	Biochemistry, Michigan State University, East Lansing, MI, 48824-1318, USA		
source	Location/Qualifiers		
	1. 4601		
	/organism="Drosophila melanogaster"		
	/organella="mitochondrion"		
	/strain="Oregon-R"		
	/db_xref="taxon:7227"		
	/dev_stage="embryo"		
	650..1022		
	/note="repeat I-A"		
	/rpt_type=tandem		
	1023..1360		
	/note="repeat I-B1"		
	/rpt_type=tandem		
	1361..1705		
	/note="repeat I-C/A"		
	/rpt_type=tandem		
	1706..2043		
	/note="repeat I-B2"		
	/rpt_type=tandem		
	2044..2388		
	/note="repeat I-C"		
	/rpt_type=tandem		
	2491..2511		
	/note="deoxythymidylate stretch"		
	2512..2648		
	/partial		
	/rpt_type=tandem		
	2649..3112		
	/note="repeat II-A"		
	/rpt_type=tandem		
	3113..3576		

OY	299	atattaaacgtagcttgacgttactccctttaaacttaataataaccagngct	358
Db	3543	ATATTAAATTTATTTAAATTAATTTATTTTTCATATATATATATATATTAATTAATAT	3484
OY	359	ttacttaacttaacttgcngatngtatacataggtttatgcncctcttlttgcncgct	418
Db	3483	CAATTAATTTATTTAAATTTTAAATATATATATTAATTAATTAATTAATTAATTAATTA	3424
OY	419	gattggtttatgcttatttgaaatgngcncagtgtaagaaatgaagaacatttatctg	478
Db	3423	AATTAAATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	3365
OY	479	gagaaataataacttaactatgctcaatatatagagaataatataatnccctgagtacg	538
Db	3364	TTTAAATTTTAAATTTTATTTTATTTAAATTTATAGATATATTAATTAATTAATTAATTTA	3305
OY	539	tatgagatcgagtagaagatctcttgaataatacttgagaacttgccttcccaaaagt	598
Db	3304	TATATAT---ATTAATATCTATTAATTTATTAATTAATTAATTAATTAATTAATTAATTA	3249
OY	599	aaatattgtagatgtaacttaagtatacacatgaaatataaaaaataataatcaaaa	658
Db	3248	AAATTAATTTTATTTTAAATTTTATTTTAAATTTGAAAAATTAATTAATTAATTTTCA	3189
OY	659	tagaaaaaactgatagtgatctacccttcaacgctttgtaactatctctggtacccccc	718
Db	3188	TATTAATTTTATTTTATTTTAAATTTTCTTTGTTTATTTTAAAAACATGATTTTATTTA	3129
OY	719	taaaccttaagttccaccaacaataaaatctcattatgcatctctatccctttaga	778
Db	3128	TAAATATTT---TTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	3072
OY	779	aagtgaacaacaatatctcaagttatatatagtttccaataaagaataaataa	838
Db	3071	TTATTTTAAATATTTTAAATTTTCTATATATATATATATTTTATTAATTAATTTCAATTA	3012
OY	839	ataaaaataatagtagttacaaaaaataaaataaataatcttaaccagcgicannaaaac	898
Db	3011	TATTAATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	2952
OY	899	actaaaacctaaacccctaaatatctaaacttttagtgaacccctaaaccttggataatc	958
Db	2951	TATATATTAATTAATTAATTAATTAATTAATTAATTTT---TATTAATTTATATCTACATTTTAAAT	2893
OY	959	ttaaacatttaaacattaaacaactaaacccctaaactccctaactctaaaccccttaagtg	1018
Db	2892	TTAAATTTTATTTTATTTTAAATTTTATGATATATTAATTAATTAATTAATTAATTTATTA	2833
OY	1019	taaatgttagtgcttttgattatagtttagagattatccaaaggtttaagtttaacc	1078
Db	2832	TAAATATCTATTAATTTAATTAATTAATTAATTAATTTTATTTTAAAAAATAATTTATTTT	2773
OY	1079	aagagtctatggtttagagatatacacttagagatttagtgcttcttagcgacgcttcaa	1138
Db	2772	TTTAAAAATTTT---TTTAAATTAATTAATTAATTAATTAATTAATTAATTTTCAATTA	2718

SOURCE	fruit fly,
ORGANISM	Mitochondrion <i>Drosophila melanogaster</i>
	Eukaryota; Metazoa; Arthropoda; Tracheata; Insecta;
	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;

REFERENCE Muscomorpha, Ephydroidea, Drosophilidae, Drosophila.
AUTHORS 1 (bases 12511 to 12682)
Clary, D.O., Goddard, J.M., Martin, S.C., Fauron, C.M. and
Wolstenholme, D.R.
TITLE Drosophila mitochondrial DNA: a novel gene order
JOURNAL Nucleic Acids Res. 10 (21), 6619-6637 (1982)
MEDLINE 83090428
REFERENCE 2 (bases 5269 to 5695).
AUTHORS Clary, D.O., Wahlthner, J.A. and Wolstenholme, D.R.
TITLE Transfer RNA genes in Drosophila mitochondrial DNA: related 5' flanking sequences and comparisons to mammalian mitochondrial RNA genes
JOURNAL Nucleic Acids Res. 11 (8), 2411-2425 (1983)
MEDLINE 83220794
REFERENCE 3 (bases 404 to 5272)
AUTHORS de Bruijn, M.H.
TITLE Drosophila melanogaster mitochondrial DNA, a novel organization and genetic code
JOURNAL Nature 304 (5923), 234-241 (1983)
MEDLINE 83245048
REFERENCE 4 (bases 804 to 1778)
AUTHORS Satta, Y., Ishiwa, H. and Chigusa, S.I.
TITLE Analysis of nucleotide substitutions of mitochondrial DNAs in Drosophila melanogaster and its sibling species
JOURNAL Mol. Biol. Evol. 4 (6), 638-650 (1987)
MEDLINE 88174373
REFERENCE 5 (bases 5268 to 13619)
AUTHORS Garesse, R.
TITLE Drosophila melanogaster mitochondrial DNA: gene organization and evolutionary considerations
JOURNAL Genetics 118 (4), 649-663 (1988)
MEDLINE 88212147
REFERENCE 6 (bases 441 to 2967)
AUTHORS Satta, Y. and Takahata, N.
TITLE Evolution of Drosophila mitochondrial DNA and the history of the melanogaster subgroup
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 87 (24), 9556-9562 (1990)
MEDLINE 91088557
REFERENCE 7 (bases 14215 to 14512)
AUTHORS Ballard, J.W., Olsen, G.J., Falt, D.P., Odgers, W.A., Rowell, D.M. and Atkinson, P.W.
TITLE Evidence from 12S ribosomal RNA sequences that onychophorans are modified arthropods
JOURNAL Science 258 (5086), 1345-1348 (1992)
MEDLINE 93088057
REFERENCE 8 (bases 14917 to 19517)
AUTHORS Lewis, D.L., Farr, C.L., Farquhar, A.L., and Kaguni, L.S.
TITLE Sequence, organization, and evolution of the A+T region of Drosophila melanogaster mitochondrial DNA
JOURNAL Mol. Biol. Evol. 11 (3), 523-538 (1994)
MEDLINE 94285822
REFERENCE 9 (bases 1 to 408, 13319 to 19517)
AUTHORS Lewis, D.L., Farr, C.L., and Kaguni, L.S.
TITLE Drosophila melanogaster mitochondrial DNA: completion of the nucleotide sequence and evolutionary comparisons
JOURNAL Insect Mol. Biol. 4 (4), 263-278 (1995)
MEDLINE 96423163
REFERENCE 10 (bases 1 to 19517)
AUTHORS Lewis, D.L., Farr, C.L., and Kaguni, L.S.
TITLE Direct Submission
JOURNAL Submitted (03-OCT-1995) Laurie S. Kaguni, Biochemistry Department, Michigan State University, East Lansing, MI 48824-1319, USA
FEATURES
source
1. 19517
/organism="Drosophila melanogaster"
/organelle="mitochondrion"
/db_xref="taxon:7227"
/note="derived from new and previously submitted sequences; sequence is a composite containing sequences obtained from different Drosophila melanogaster strains"
1. 65
/product="tRNA-tle"
complement(97, 165)
tRNA
tRNA

tRNA
CDS
/product="tRNA-Gln"
171..239
/product="tRNA-Phe"
240..1265
/codon_start=1
/transl_table=5
/product="NADH dehydrogenase subunit 2"
/protein_id="AAC47811.1"
/db_xref="GI:1166530"
/translation="MRNNSKILFITMIIGILITVTSNWGAWGLEINLSFPL
LSNNNMSTESALKIFLIDVLASVLFSSILMLKNNKNNNEINSEFTSMISML
LKSAPFHEWFEPMMEGLIMANALMIMWOKTAPMLISYINIKILISVLSYII
GAIGLNQTSIRKLMAFSSINLGMALSSIMSESIWLIIFPYSPLSPVLFPMFIF
KLPHNLQFSWFNSKIKFTLPEMLISGGLPPLGPIPKVILQOOLICNOYRMT
IMMSLTILIFFLYRICYSAFMNMYENNWMKMNNSINYMYIMTFSTIFGLFLI
SLFYPFW"
1264..1329
/product="tRNA-TTP"
complement(1322..1383)
/product="tRNA-Cys"
complement(1403..1468)
/product="tRNA-Tyr"
join(1470..1472,1474..3009)
/note="mechanism underlying reading frame shift after first codon uncertain"
/codon_start=1
/transl_table=5
/product="cytochrome c oxidase subunit I"
/protein_id="AAC47812.2"
/db_xref="GI:7412849"
/translation="MSKWLSTNKHKIDGTYLIFGAWAGVGTSLIRAEIHPG
ALIGDQIYNIVTAHAFIMFEWPMITIGFGMWLVPIMGAPDMAPPMNNSEF
LILPALSLIVSSWENGAGTCWVYPPSLAGISAVDLAIFSLAGISSTIGA
VNEITTVINMRSTGISLDRMPLEWVSVITALLLSLIPVAGATMTLDRNITSF
FPAGGDEPILKOLHFEFGPEEYVILIPGCMTSHIISGSKKEFGSGIYAM
LAIGLIGETVMAHMEFTGMVDTPAYFTSAMITAVTGKITSWALTIGTOLST
PALMAIGFVPLFTYGLGGLAVLASSVDIILHDITYVAHFHYLSGAVFAIMAG
IHMYPLFTGLTNNKWLKSHFTIMFTGVNLTFEPQHFGLAGMPRRSDYDAVTNN
IVTIGTSTLSLGLIFFFTIIESLVSOROVITPQLNSSITEMYQMPARHSSELP
LITN"
3012..3077
/product="tRNA-Leu"
3083..3767
/note="TAA stop codon is completed by the addition of 3' A residues to the mRNA"
/codon_start=1
/transl_except=(pos:3767,aa:TERM)
/product="cytochrome c oxidase subunit II"
/protein_id="AAC47813.1"
/db_xref="GI:1166532"
/translation="MSTVANGLDSDASPLMEQIIFPHDALLVMTAVLYLMEM
LFENYVNRFLHQLIEMITLIPALIFLALPSRLLYLDELINSEPVTKSIH
QWYSVEYSDNNNEIPDSYMPITMELMTDSEFRILDVDRVAVLPNNSQIRLVAAVY
HSWVPAIGVAVDGTPLRNLQTNFIRNPGLFYGQCEICGAMHSEMPVIESPVNY
FIKWSNNS"
3768..3838
/product="tRNA-Lys"
3840..3906
/product="tRNA-Asp"
3907..4068
/codon_start=1
/transl_table=5
/product="ATPase 8"
/protein_id="AAC47814.1"
/db_xref="GI:1166533"
/translation="MPQMAPISWLLFIIFISITFLFCSINYSYMPNSPKSNEIKNI
NLSNMNKKW"
4062..4736
/codon_start=1
/transl_table=5
/product="ATPase 6"
/protein_id="AAC47815.1"

RESULT	11
LOCUS	AC005504/c
DEFINITION	104992 bp DNA linear HTG 01-APR-1999 ***, 3 unordered pieces.
ACCESSION	AC005504
VERSION	AC005504.3 GI:4558584
KEYWORDS	HTG; HTGS_PHASE1.
SOURCE	malaria parasite P. falciparum. Plasmodium falciparum
ORGANISM	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE	1 (bases 1 to 104992) Hyman,R.W., Fung,E.L., Qin,F., Tamaki,T., Kurdi,O.B., Conway,A.B. and Davis,R.W. Plasmodium falciparum 3D7 chromosome 12 Unpublished
TITLE	2 (bases 1 to 104992)
REFERENCE	Hyman,R.W., Qin,F., Fung,E.L., Conway,A.B. and Davis,R.W.
AUTHORS	Direct Submission
TITLE	Submitted (21-AUG-1998) Stanford DNA Sequencing and Technology
JOURNAL	

center, Stanford University, 855 California Avenue, Palo Alto, CA
94304, USA

On Apr 2, 1999 this sequence version replaced gi:4337172.

NOTE: This is a 'working draft' sequence. It currently consists of 3 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. * This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

*
1 58642: contig of 58642 bp in length
* 58643 gap of unknown length
* 58843 contig of 32169 bp in length
* 91012 gap of unknown length
* 91212 contig of 13781 bp in length.
Location/Qualifiers
 1..104992
organism="Plasmodium falciparum"
db_xref="taxon:5833"
chromosome="12"

BASE COUNT 44286 a 9326 c 9584 g 4141 t 405 others

ORIGIN

Query Match 6.0%; Score 171.4; DB Z: Length 104992;
Best Local Similarity 47.0%; Pred. No. 1.4e-14;

Matches 725; Conservative 0; Mismatches 803; Indels 13; Gaps 6

OY 468 attatatttgagagaataataattaactaatgctcaatatatagagaaaattatnccct 527
DB 74679 ATTAAATTAATAATTAAATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 74620
OY 528 tgagtcatcgtacgatcgagtagaaagatccttgaataatattgagaaccttccttt 587
DB 74619 TATTATTAATAAATNAATTTAAATTAATTAATTAATTTTAAATTAATTAATTAATTAANF 74560
OY 588 ttccaanaagta----aacattggatgytaactaaqtaacaatgaaaaattaaan 643
DB 74559 AAAATTAATAAATAATTAATTAATTTTAAATTAATAATTAATTAATTAATTAATTAATTA 74500
OY 644 aaatttaaaccataagagaaaacgcgaaggagatccccctaacttgtgaacttat 703
DB 74499 TTAAATTAATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATAAAAATA 74440
OY 704 tccttggtcaccccctaaccctaagctcaagctaccaacaacataatcatactgcatat 763
DB 74439 AATTAATTAATCAATTTAAATTTATTAATTAATTAATTTAAATTAATTAATTAATTAATNAT 74381
OY 764 tctatactctttaagaaaguguaaccaaataatatcataagttatatatglttccaataa 823
DB 74380 ATTTAAATTAATTTAAATTAATTAATTAATTAACCTTTAAATTAATTAATTAATTAATTAAT 74321
OY 824 aagaataaanaataaalaa--aaataatagagtlccaaaaaanaaataatfatllt 881
DB 74320 ATAATTAATTAATTAATAATTTAAATTCATATATTTTTTAATTAATAATAATTTCTTAATTTATT 74261
OY 882 taacagcgctcanaaacaactaaaccctaaccctaataliaaactttlagttaaaccot 941
DB 74260 TTATTACAATTAATTAATAATTAATTAATTTTAAATTAATTAATTAATTAATTAATTAATTTAT 74201
OY 942 aaaccttggagaanaactttaaacatlaaacatlaaacacaacccaacccaactcctaaact 1001
DB 74200 TAATTAATTAATTAATTAATTAAGTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 74141
OY 1002 ctaaaccttaagtgltttaaagtgtagtggtttggattatagttttagggattacoaa 1061
DB 74140 AAATTAATTAATTAATTAATTAATTAATTTTATTATATATATATATATATTAATTAATTAAT 74081
OY 1062 aggtttaaggttacccaagagttatggttttaggagtaagactaggattagtggtt 1121
DB 74080 TAAATTAATAATTAATTAATTAAGAAGATTAATTTTACTTTTAAATTAATTAATTAATTAATG 74021

QY 1122 tactgagcgtcacaagatatttttaaaaaatatttttctgtaacaactactatttt 1181
 Db 74020 AATTAATTAATTTATGTTATTTATTAATTAATTTATTTATTTATTTATTTAA 73961
 QY 1182 attatttttttaaccttttatataataaataataataatttaaccatctgttt 1241
 Db 73960 TAAATTAATTAATTTATTAATTAATTTATTTATTTATTTATTTATTTATTTAA 73901
 QY 1242 atataagtcgtcgttaacatttttttctgtaacaataaagtcgaacttagattc 1301
 Db 73900 TATTTAATTAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAA 73841
 QY 1302 caatgcaaa-attatttttttctgtaacaataaattatgcaagtcgtcgttata 1360
 Db 73840 AATTAATTAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAA 73781
 QY 1361 aataatttatactcacaagtcgtatgtgcaacaatgtaacatgtaataagaaacta 1420
 Db 73780 GTGTAATTAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAA 73721
 QY 1421 attatt 1480
 Db 73720 TATCTTAATTAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAA 73661
 QY 1481 aacgaattgagtaatt 1540
 Db 73660 TATGTAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAA 73601
 QY 1541 atttttctgtaagtcgaactcaaaatcccaagaagtcgaacgcaagaataagtc 1600
 Db 73600 TTTAATTAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAA 73541
 QY 1601 acgtttgaggaagaaactgtaacatgagcccaagtcgaacaaactcgttatttca 1660
 Db 73540 ATTATTTGTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAA 73481
 QY 1661 aatccttaacaattttagagagtaagtaagcaacaacacatgtaacatcatagaa 1720
 Db 73480 AATTAATTAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAA 73421
 QY 1721 tattccttaaaatttactagcgaataataaacttttttttttttttttttttttttttttt 1780
 Db 73420 TATTTAATTAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAA 73361
 QY 1781 aaaaacttttataataactaaacatttttaataaagaataaaggaacacatgagat 1840
 Db 73362 TTTATTTAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAA 73303
 QY 1841 acaataaataatagttatttcttgaagatagataataataataaccagtcacat 1900
 Db 73302 TTTATTTAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAA 73244
 QY 1901 atttacaataataataataattttcgttcgtaacatcatcattacataataatcataaa 1960
 Db 73245 AATTTATTTAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAA 73186
 QY 1961 ccaatgtagatgatactttatttattatataataataa 2001
 Db 73185 TTTATTTAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAA 73145

RESULT 12
 AC004157/c 169546 bp DNA linear HTG 12-AUG-2000
 LOCUS Plasmodium falciparum chromosome 12 clone 3D7, *** SEQUENCING IN
 DEFINITION PROGRESS ***, 2 unordered pieces.
 AC004157
 AC004157 GI:9797712
 VERSION HTG: HTGS_PHASE1
 KEYWORDS malaria parasite P. falciparum.
 SOURCE Plasmodium falciparum
 ORGANISM Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 REFERENCE 1 (bases 1 to 169546)

AUTHORS Hyman, R.W., Fung, E.L., Qin, F., Rowley, D., Mao, J., Tamaki, T.,
 TITLE Kurdi, O.B., Conway, A.B. and Davis, R.W.
 JOURNAL Plasmodium falciparum 3D7 chromosome 12
 REFERENCE Unpublished
 AUTHORS 2 (bases 1 to 169546)
 TITLE Hyman, R.W., Qin, F., Fung, E.L., Conway, A.B. and Davis, R.W.
 JOURNAL Direct Submission
 COMMENT Submitted (19-FEB-1998) Stanford DNA Sequencing and Technology
 94304, USA
 On Aug 12, 2000 this sequence version replaced gi:8810447.
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 2 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 * 1 23466: contig of 23466 bp in length
 * 23467 23666: gap of unknown length
 * 23667 169546: contig of 145880 bp in length.
 * Location/Qualifiers
 source
 1. 169546
 /organism="Plasmodium falciparum"
 /db_xref="taxon:5833"
 /chromosome="12"
 /clone="PFYAC293"
 /clone="3D7"
 BASE COUNT 69871 a 15381 c 15705 g 68389 t 200 others
 ORIGIN

Query Match 5.0%; Score 171.4; DB 2; Length 169546;
 Best Local Similarity 47.0%; Pred. No. 1,3e-14;
 Matches 725; Conservative 0; Mismatches 803; Indels 13; Gaps 6;

QY 468 atttatacttgaggaataataatatttctgtaacatgagagaataattatccct 527
 Db 86263 ATTATTAATTAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAA 86204
 QY 528 tgaacttactgtaagtcgagtagaagaactcttgtaataatttggagaactgcctt 587
 Db 86203 TATTTAATTAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAA 86144
 QY 588 tcccaaaagtaa---aatattgtagtaacttaagtaacatgaaacttaaaaa 643
 Db 86143 AAAATTAATTAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAA 86084
 QY 644 aaaaattcaaatcagaataaactgtagtgcacccctcaacgctttagaacttat 703
 Db 86083 TTTATTTAATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 86024
 QY 704 tcttggttccccccttaaaccttaagttccacaacaataaatttcatattgcatat 763
 Db 86023 AATTAATTAATTTAATTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAAT 85965
 QY 764 tctatattcttttgaagaagtgaaacaaataatttcaagttatatttatttccaataa 823
 Db 85964 ATTAAATTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAA 85905
 QY 824 aaagataaaaaaataaa---aaataatagtgatcacaataaaaaaataatattt 881
 Db 85904 AATTAATTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAA 85845
 QY 882 taccagcgtcanaaaacacaaacctaacaacctaataattttaggtaaacctt 941
 Db 85844 TTTATTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAA 85785
 QY 942 aaaccttggataaacttataacatttaaaacataaaacataacctaataacctaact 1001
 Db 85784 TATTAATTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAA 85725

OY 1002 ctaaaccccttaagtggttaaatgttggttgattagtaggttagatccaa 1061
 DB 85724 ATATTATTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 85665
 OY 1062 aggttaaggttaccagaaggttaggttaggttaggttaggttaggtt 1121
 DB 85664 TAATTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 85605
 OY 1122 tactgagcagcttaagtagtttaaaatatttttttgtaacacacatttt 1181
 DB 85604 AATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 85545
 OY 1182 attattttttacccttttatataaacaataataacttaactcactcgttc 1241
 DB 85544 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 85485
 OY 1242 atattaaagtcattgtaacatttttttggttaacaaaatgtcacttagaattc 1301
 DB 85484 TAATTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 85425
 OY 1302 caatgcacaa-attattattttcagctaaatgaatgaagtgatgacttata 1360
 DB 85424 AATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 85365
 OY 1361 aataatttattatctcaaatgctataatgccaacatgtaactaagaactta 1420
 DB 85364 GTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 85305
 OY 1421 attattattcattatttttttttttttttttttttttttttttttttt 1480
 DB 85304 TATCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 85245
 OY 1481 aaacgaattgagtaataatgttcttctttaaagaatgcgaatacgaatacaca 1540
 DB 85244 TTATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 85185
 OY 1541 atttattgtctgtagaaccataaatcatcctcaagaagtgagcagaagaatga 1600
 DB 85184 TTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 85125
 OY 1601 acgttttgagggaagcctaactgccccagaagtcataaactcttcttgatttaca 1660
 DB 85124 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 85065
 OY 1661 aatccctacaatttagtagagtagtaagacaacacatgattatcatcatatagaa 1720
 DB 85064 ATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 85005
 OY 1721 tattctaaaaaattaccagcgaataaataactcttcttatttatacaaatccctat 1780
 DB 85004 TAATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 84947
 OY 1781 aaaaactatttatataactaaacattttaaaagaagaataagggaccatgatt 1840
 DB 84946 TTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 84887
 OY 1841 acataaataatagttattcttaagaatgataaataataataataacagctccat 1900
 DB 84886 TTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 84828
 OY 1901 attatcaaaaataaataacttttgcgtacgcgaataatcattactataaatacaaaa 1960
 DB 84829 ATTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 84770
 OY 1961 ccaatgtagatgataattttatttatataatataaataa 2001
 DB 84769 TTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 84729

RESULT 13
 AC093899 172816 bp DNA linear PRI 05-FEB-2002
 LOCUS AC093899
 DEFINITION Homo sapiens chromosome 2 clone RP11-724016, complete sequence.

ACCESSION AC093899 AC068884
 VERSION AC093899.3 GI:18497265
 KEYWORDS HTG.
 SOURCE
 ORGANISM Homo sapiens
 human.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE
 AUTHORS Waterston,R.H.
 TITLE The sequence of Homo sapiens clone
 JOURNAL Unpublished
 REFERENCE
 AUTHORS Waterston,R.H.
 TITLE Direct Submission
 JOURNAL Submitted (10-SEP-2001) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 REFERENCE
 AUTHORS Waterston,R.H.
 TITLE Direct Submission
 JOURNAL Submitted (05-FEB-2002) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 COMMENT On Feb 5, 2002 this sequence version replaced gi:15625013.
 FEATURES
 source Location/Qualifiers
 1..172816
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="2"
 /clone="RP11-724016"
 BASE COUNT 53405 a 33788 c 33104 g 52519 t
 ORIGIN
 Query Match 5.9%; Score 169.6; DB 9; Length 172816;
 Best Local Similarity 46.9%; Pred. No. 2.2e-14;
 Matches 762; Conservative 0; Mismatches 840; Indels 22; Gaps 7;
 OY 382 ttatcatggttaatgcnctcttttttgancocgttgattgtaacttgaatt 441
 DB 22715 TGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 22774
 OY 442 gtlgcnacgtaagaatgaagaacattatatttgagagaataataatataatglt 501
 DB 22775 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 22834
 OY 502 caatatatgagaataatataatccttgatgctgtagtgagtagagagatcct 561
 DB 22835 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 22894
 OY 562 tgaataatattggaactgcttctccaaagaataaataattgtaactta 621
 DB 22895 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 22950
 OY 622 gttacacatgagaatlaaaaaaataaatacaaatagaanaaacctgtagtacta 681
 DB 22951 TAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 23010
 OY 682 cccctcaacgtttgaaactatcttggttcaaccccccaaaccttaagtgaccaca 741
 DB 23011 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 23070
 OY 742 ataaaattcattatgataatcattttagaagaagaacaaaataattatacaag 801

TNKTKYKRSKQSLFEDGKLNTHDKLILFEKRLPYKCKAYCEISAKYEVYKYLIDDEKCK
 ENYISLIDCIIOSVKLIDFELDCFPDYNFYEYKINDKVLKINLEIFERKNDIYFHR
 ELICKCKIMSYIHENANGNELIPELLYFEFWKNDKMLILFVNYFVNYVDHLYFLRN
 HEIKYKLETFKRYKLNNSNIPFNKRLIOEMEFNLITYPELKEINENRYIKKMKELIYKCK
 CFAFHEWVDHIDMKELINTLRLAYVDSLIDIDINMKKACIMNNLILENTEYISKIL
 NFYCTLIKRKYDMDITYKLEKVIKATHHILCKDCKTYKLEFCGDIYDSTLNSKMRK
 FLINKIDKNEILFEYECILKILKRIKVENFOSCILSLISLILNRYIYVNNV
 LEFNDIMKSYLCMIFLCKRIKTEENAVLILHNDNDNENKREKINDIIOKRIKEY
 IFEYMEKNTCKDHEFKLSDILSLISTLNSNTPKILEYVNSDYPFLIPNNSICLILFV
 NRSVKKRYKDPYIYIINDISPRYKITYKNDRTKRRKKEVFLSSMKELICNLILSVN
 RYIKHIEHEDNEFDOKDOYVCSILFNLNIPFKRLIHEHYTNLMCHVYKTYVFNCKML
 INEDILSLITLDCSKFOFLENNSDNRCRELILHLYNIIDILIKNYLNTKYSIDT
 NISYFISLSKSCYKVENMLILLESLOSEKPEYKNSCKSGHIGMOONLIDNNSCE
 KISYERLYEKKENYFENLKNKILIECLILIEYLYKRYKTYLAIKOSCPINLENTL
 KILYILANNLYFEMYGVCYEMLEFRYSIKHEOLFCYVNNKNEHEMKPCTIICHISE
 DDYIEMSNNTMYVLFEYDKINNSBROSNILRNSTNDRFLIDEIKERYKLANNTLILH
 NNVDLNEKSNNSNGSNLILKDKKNNHNNVENDLIDNKNRKLIDOKOGNGBEN
 CKDVLANDIINIBGFLMKKEKKFLFEDJILNFKLITFRKRYVSSSLFIMADVEKILIK
 DMNKIYICLEMYKIKNECEKAVLYIYIDILFKERK

```

Query Match      5.9%; Score 168.4; DB 3; Length 14867;
Best Local Similarity 45.5%; Pled. No. 5,2e-14;
Matches 715; Conservative 0; Mismatches 848; Indels 9; Gaps 3;

QY 469 ttatatltggagagaaatataatlaattlaaaltgltcaatatataatgagaaatatiaaccct 528
DB 8106 TATTTAAATGAATAAATAAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 8165
QY 529 gatgttactgtaatgagtgagagagaagcttcttgataatatatttgagaacttgcccttt 588
DB 8166 TATAATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 8225
QY 589 ctcaaaaatlaaaaataattgataatgaactaaatgaatgaacactggaactaaataaaaaaat 648
DB 8226 TTTTAAAAAAAANANVTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 8285
QY 649 taaataaaaatgaaaaaactgactgtaatcacaaccttaacgcttttgaaacttacttg 708
DB 8286 TAATTTAAATTTAAATTAATTAATTAATTAATTTTAAATTAATTAATTAATTAATTAATTAAT 8345
QY 709 gtacaccccttaaacctcctaagtlcacaacaaataaaattccattatgcacttcat 768
DB 8346 AGTCATTTATTTATTTATTTTATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 8405
QY 769 alcctttgaaagtgaaacaaataattacaagttaattalagtlttccaataaaaaaga 828
DB 8406 TTGTTTCTTTGTAATTAATTAATTAATTAATTAATTTTAAATTAATTAATTAATTAATTAAT 8465
QY 829 taaaaaatlaaaataaaataatagtagttcaacaaataaaaaaattaaac-----atttta 883
DB 8466 TTAATAAATAAATTAATTAATTAATTAATTCATTTATTTTAAATTAATTAATTAATTAATTAAT 8525
QY 884 ccaagctcanaanaacactcaaaaacctaaacctaaatlaaacittlttagtfaaacoccaa 943
DB 8526 TTTCCAAATAATAAATAAATAAAGTATTTTAAAGTTATTAATAATTAATTAATTAATTTTAA 8585
QY 944 acccttggataaactttaaacattbaaacattaaaaccttaaaccttaaaccttaaacctc 1003
DB 8586 TAAATTAATAATATGCTTTAA-TTCTTAATTAATAATAATTTTAAATTAATTTTAAATAATA 8644
QY 1004 aaaccttaagtglttaaatgtttagtglttttgatattatgcttttagattatccaaga 1063
DB 8645 ATTATGATATATAATTAATTAATTAATTAATAATTAATTAATTAATTAATTAATTAATTAAT 8704
QY 1064 gtttaagtttaccacaagaagtttatgcttttgaggataatgacttaagattagtglttla 1123
DB 8705 ACATATTAATTTACACAAATAATTTATTTTAAATTAATTAATTTTAACTAATTTTAAAT 8764
QY 1124 ctgacgagcttcaaaagtaattttttaaataaatttttttgtaaacactcctaattttat 1183

```

[illegible]

LOCUS	AC0093899/c	AC0938899	172816 bp	DNA	linear	PRI 05-FEB-2002
DEFINITION		Homo sapiens chromosome 2 clone RP11-724016, complete sequence.				
ACCESSION		AC0938899	AC0068884			
VERSION		AC0938899.3	GI:18497265			
KEYWORDS		HTG.				
SOURCE		human.				
ORGANISM		Homo sapiens				
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				

Mon Jul 1 08:32:46 2002

QY 1988 atatatatataa 2002
|||
Db 22903 TATTAATATAAGA 22889

Search completed: June 28, 2002, 13:44:36
Job time: 8293 sec

us-09-509-188-3.rge

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 28, 2002, 11:40:33 ; Search time 308.16 Seconds
(without alignments)
15895.499 Million cell updates/sec

Title: US-09-509-188-3

Perfect score: 2853
Sequence: 1 ggaaccacaaagaacacg.....ccgtcttgatttaactt 2853

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 segs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

N.Geneseq_032802:*

- 1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
- 2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
- 3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
- 4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
- 5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
- 6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
- 7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
- 8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
- 9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
- 10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
- 11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
- 12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
- 13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
- 14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
- 15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
- 16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
- 17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
- 18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
- 19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
- 20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
- 21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
- 22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
- 23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
- 24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2844	99.7	2853	20	AAAX26842
2	462	16.2	674	20	AAAX26841
3	393.2	13.8	497	20	AAAX26840
4	196.4	6.9	730	21	AAAC34997
5	184.4	6.5	461	21	AAAC48164
6	174.8	6.1	2853	20	AAAX26842
7	123.6	4.3	6641	24	ABL32315
8	122.4	4.3	15548	24	ABL34155
9	121	4.2	5930	24	ABL32517

C 10	120	4.2	3795	22	AAAC87646
C 11	120	4.2	9965	24	ABL33527
C 12	119	4.2	6175	24	ABL33307
C 13	118.8	4.2	18154	24	ABL33254
C 14	118.6	4.2	6681	24	ABL32155
C 15	118.6	4.2	13548	24	ABL34155
C 16	118.4	4.2	18585	24	ABL34609
C 17	117.6	4.1	5413	22	AAAS46694
C 18	116.8	4.1	6881	24	ABL33381
C 19	116.4	4.1	5647	24	ABL33566
C 20	116.4	4.1	5647	24	AAAS61320
C 21	114.8	4.0	17183	24	ABL32486
C 22	114.6	4.0	38342	22	AAAS46746
C 23	114.4	4.0	7167	24	ABL32400
C 24	114.2	4.0	6244	24	ABL32484
C 25	113.8	4.0	18683	24	ABL33213
C 26	113	4.0	11691	24	ABL34240
C 27	113	4.0	12393	24	ABL33263
C 28	112.6	3.9	6106	22	AAAS46429
C 29	112.6	3.9	6106	24	ABL33472
C 30	111.8	3.9	3231	24	ABL34288
C 31	111.6	3.9	18218	24	ABL33949
C 32	111.2	3.9	14095	24	ABL32477
C 33	110.8	3.9	6127	24	ABL33614
C 34	110.8	3.9	15732	22	AAAS45388
C 35	110.6	3.9	16287	24	ABL32673
C 36	110.2	3.9	5163	24	ABL33249
C 37	110	3.9	6092	24	AAAS61080
C 38	109.8	3.8	6182	24	ABL34015
C 39	109.8	3.8	13574	24	ABL33317
C 40	109.8	3.8	16287	24	ABL32673
C 41	109.6	3.8	17211	24	ABL32654
C 42	109.4	3.8	1671	13	AAO24134
C 43	109.2	3.8	18624	22	ABL33702
C 44	109	3.8	15732	24	AAAS45389
C 45	108.8	3.8	5979	22	AAAS45313

ALIGNMENTS

RESULT 1	
AAAX26842	standard; cDNA: 2853 BP.
XX	XX
XX	AAAX26842:
AC	22-JUN-1999 (first entry)
XX	XX
DT	22-JUN-1999 (first entry)
XX	XX
DE	-Microspore-specific plant promoter cDNA clone Bm3.4.
XX	XX
KW	Microspore-specific plant promoter. Brassica napus cv. Brutor;
KW	gametophytic male sterility; inducible fertility; ss.
XX	XX
OS	Brassica napus.
XX	XX
PN	FR2768745-A1.
XX	XX
PD	26-MAR-1999.
XX	XX
PF	23-SEP-1997; 97FR-0011812.
XX	XX
PR	23-SEP-1997; 97FR-0011812.
XX	XX
PA	(INRG) INRA INST NAT RECH AGRONOME.
XX	XX
PI	Drouaud J, Fourgoux A, Guerche P, Pelletier G;
XX	XX
DR	WPI, 1999-217496/19.
XX	XX
PT	Microspore-specific plant promoter from Brassica napus - and vectors
PT	for producing plants exhibiting gametophytic male sterility with
PT	inducible fertility

XX Claim 1, Page 15, 32pp; French.

PS The present sequence represents a clone of a microspore-specific plant
CC promoter from *Brassica napus* cv. Brutor. The promoter is used in a
CC method for producing plants exhibiting gametophytic male sterility
CC with inducible fertility.

XX Sequence 2853 BP; 961 A; 506 C; 381 G; 996 T; 9 other.

Query Match 99.7%; Score 2844; DB 20; Length 2853;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2853; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ggaaccacaagaagaacgaagaagcaagtttcccaacttcaataataattgt 60
DB 1 ggaaccacaagaagaacgaagaagcaagtttcccaacttcaataataattgt 60
QY 61 ttcaagccatcaatgtacaacaactccttagctcaatgtgtataaagtgtgttag 120
DB 61 ttcaagccatcaatgtacaacaactccttagctcaatgtgtataaagtgtgttag 120
QY 121 attcaataaaccgggttcgagtcataagacttgaacatttccacattttaaaagt 180
DB 121 attcaataaaccgggttcgagtcataagacttgaacatttccacattttaaaagt 180
QY 181 aacgcacatcgcgtgacgttcgcatacagaagtgaacactgtctatataatgtag 240
DB 181 aacgcacatcgcgtgacgttcgcatacagaagtgaacactgtctatataatgtag 240
QY 241 atttaaaagtgaacacacgatacgcgtgacgttcgcatacagaagtgaacactgt 300
DB 241 atttaaaagtgaacacacgatacgcgtgacgttcgcatacagaagtgaacactgt 300
QY 301 atttaacgtatgttgaacgttatccttlttaaatcttaataataacagngcttt 360
DB 301 atttaacgtatgttgaacgttatccttlttaaatcttaataataacagngcttt 360
QY 361 acttataaatttngcatngttatcagttttagtgcnccttttttttgganccgttga 420
DB 361 acttataaatttngcatngttatcagttttagtgcnccttttttttgganccgttga 420
QY 421 ttgtttatgcttatttgaatgtngcmnagtaagaataagaacaattatatttga 480
DB 421 ttgtttatgcttatttgaatgtngcmnagtaagaataagaacaattatatttga 480
QY 481 gaaataataatttaatttcaatatatagagaataataatcttctgtatgtactgta 540
DB 481 gaaataataatttaatttcaatatatagagaataataatcttctgtatgtactgta 540
QY 541 tggatgcgagtagaagaatcttgaataataattgaagactgtcctttcccaaaagt 600
DB 541 tggatgcgagtagaagaatcttgaataataattgaagactgtcctttcccaaaagt 600
QY 601 aatatttgaatgttaacttaagttacaacatgaaataataaaaaaattaaacaaata 660
DB 601 aatatttgaatgttaacttaagttacaacatgaaataataaaaaaattaaacaaata 660
QY 661 gaaaaaacgtatgttgaaccccttcaacgttttgaacttatttggttaccacctta 720
DB 661 gaaaaaacgtatgttgaaccccttcaacgttttgaacttatttggttaccacctta 720
QY 721 aaccttaagtltacccaacaataaatttcatattgtcatattctataatcttttga 780
DB 721 aaccttaagtltacccaacaataaatttcatattgtcatattctataatcttttga 780
QY 781 gtgaacaataattatcaagttatattgttttcaataaagaataaanaaataat 840
DB 781 gtgaacaataattatcaagttatattgttttcaataaagaataaanaaataat 840
QY 841 aaaaaataatagtttacaataaaaaaataatatttccgcggtcnaaaacac 900
DB 841 aaaaaataatagtttacaataaaaaaataatatttccgcggtcnaaaacac 900

DB 841 aaaaaataatagtttacaataaaaaaataatatttccgcggtcnaaaacac 900
QY 901 taaaacctaaacccataatatttaaacctttagtgaacccataaccccttgaataatct 960
DB 901 taaaacctaaacccataatatttaaacctttagtgaacccataaccccttgaataatct 960
QY 961 aaacattaaacattaaacacataaacccataacccataacccataaccccttaagt 1020
DB 961 aaacattaaacattaaacacataaacccataacccataacccataaccccttaagt 1020
QY 1021 aatgttagtgtttttagttatagtttaggtattcccaagtttaagtttaccaca 1080
DB 1021 aatgttagtgtttttagttatagtttaggtattcccaagtttaagtttaccaca 1080
QY 1081 gaatttagttagttaggtatttagttaggttaggttttctgaacaggttcaagt 1140
DB 1081 gaatttagttagttaggtatttagttaggttaggttttctgaacaggttcaagt 1140
QY 1141 atttttaaaaaatatttttttgaacacatactattttatttttttacccttt 1200
DB 1141 atttttaaaaaatatttttttgaacacatactattttatttttttacccttt 1200
QY 1201 tataataaacaataataataatttaatactcctgtttcatatgaatgtcatgttaa 1260
DB 1201 tataataaacaataataataatttaatactcctgtttcatatgaatgtcatgttaa 1260
QY 1261 catatttttttgaacaaaaaatgtcaactttagaattccaatgtcaaatatttat 1320
DB 1261 catatttttttgaacaaaaaatgtcaactttagaattccaatgtcaaatatttat 1320
QY 1321 ttcttcgctaaatttaattgcaaggtcattttagcttataataaatttattatctca 1380
DB 1321 ttcttcgctaaatttaattgcaaggtcattttagcttataataaatttattatctca 1380
QY 1381 atgctatattgttgcacaacatgtgttaataagaacataattatatttattatctca 1440
DB 1381 atgctatattgttgcacaacatgtgttaataagaacataattatatttattatctca 1440
QY 1441 tcttaatcgtgtgaacaaatgtcaaaatattatttgaacgaatgtgaatatt 1500
DB 1441 tcttaatcgtgtgaacaaatgtcaaaatattatttgaacgaatgtgaatatt 1500
QY 1501 tgttcaattttttaaagaatcgaatgaataaacaacaaatttattttagtgaatgaac 1560
DB 1501 tgttcaattttttaaagaatcgaatgaataaacaacaaatttattttagtgaatgaac 1560
QY 1561 taaaatcctatccttaagaaggtgaacgcaagaataagtcacagtttggggaagctaac 1620
DB 1561 taaaatcctatccttaagaaggtgaacgcaagaataagtcacagtttggggaagctaac 1620
QY 1621 tatggcccaaggtcacaacatcttctgtattatcaaaatccttacaatttagt 1680
DB 1621 tatggcccaaggtcacaacatcttctgtattatcaaaatccttacaatttagt 1680
QY 1681 gatttaatgacaaacacatgattatcatcattatgaatattttaaanaatttagtgc 1740
DB 1681 gatttaatgacaaacacatgattatcatcattatgaatattttaaanaatttagtgc 1740
QY 1741 gaataattaaaatcttcttattatcacaatccttataaanaacttattatataac 1800
DB 1741 gaataattaaaatcttcttattatcacaatccttataaanaacttattatataac 1800
QY 1801 taaaacaatttaattaaagaanaaataagggcattgatatataatattgtatt 1860
DB 1801 taaaacaatttaattaaagaanaaataagggcattgatatataatattgtatt 1860
QY 1861 tcttaagatagttatataatataacaggtccatatttttcaaaaataataa 1920
DB 1861 tcttaagatagttatataatataacaggtccatatttttcaaaaataataa 1920
QY 1921 ttcttcgtagtcgtaatacattataataattcaaaaacacatgtgatatatt 1980
DB 1921 ttcttcgtagtcgtaatacattataataattcaaaaacacatgtgatatatt 1980

QY 1981 tattatataataataaaccctaagccttaccactcgataaccataaactttt 2040
DB 1981 tattatataataataaaccctaagccttaccactcgataaccataaactttt 2040
QY 2041 ctctcgttcgtctaactcaaggtctcgaaagtaaaaaaataatgaagatgtcacac 2100
DB 2041 ctctcgttcgtctaactcaaggtctcgaaagtaaaaaaataatgaagatgtcacac 2100
QY 2101 ttgttcgtctatgataccttcttctaagctgtgtacatcccaagttacagcaacgaac 2160
DB 2101 ttgttcgtctatgataccttcttctaagctgtgtacatcccaagttacagcaacgaac 2160
QY 2161 tagagatcaactaactaaccagaagccttctctatcgcgcaacttcttcacccca 2220
DB 2161 tagagatcaactaactaaccagaagccttctctatcgcgcaacttcttcacccca 2220
QY 2221 agcaacattgacggtctcggtgctccgaaagccttccctcgagccacttcgactaa 2280
DB 2221 agcaacattgacggtctcggtgctccgaaagccttccctcgagccacttcgactaa 2280
QY 2281 ctccgttccatctgacacagaagatgtaacagatgtgagaagcaagaagatgta 2340
DB 2281 ctccgttccatctgacacagaagatgtaacagatgtgagaagcaagaagatgta 2340
QY 2341 catgttttaatatgatacgtctgagacttcttccacagaagcgttatgtgagtggaat 2400
DB 2341 catgttttaatatgatacgtctgagacttcttccacagaagcgttatgtgagtggaat 2400
QY 2401 gttgcgcgcgacatcaagaagatgtaacagatgtgagaagcaagcgttcttgccttcc 2460
DB 2401 gttgcgcgcgacatcaagaagatgtaacagatgtgagaagcaagcgttcttgccttcc 2460
QY 2461 atgaccccttcttgaccggtctatgtaagctacatgtctccacggttgttgatctact 2520
DB 2461 atgaccccttcttgaccggtctatgtaagctacatgtctccacggttgttgatctact 2520
QY 2521 caactctctcttcaacagagctccttcaatgctcttcttcaacagagctccttcaatgctc 2580
DB 2521 caactctctcttcaacagagctccttcaatgctcttcttcaacagagctccttcaatgctc 2580
QY 2581 ctctacatgctccttcaacagagctccttcaatgctccttcaatgctccttcaatgctc 2640
DB 2581 ctctacatgctccttcaacagagctccttcaatgctccttcaatgctccttcaatgctc 2640
QY 2641 ctctacatgctccttcaacagagctccttcaatgctccttcaatgctccttcaatgctc 2700
DB 2641 ctctacatgctccttcaacagagctccttcaatgctccttcaatgctccttcaatgctc 2700
QY 2701 ctctacatgctccttcaacagagctccttcaatgctccttcaatgctccttcaatgctc 2760
DB 2701 ctctacatgctccttcaacagagctccttcaatgctccttcaatgctccttcaatgctc 2760
QY 2761 tattcaagatgatagtcctcaggaagcttcttcttcttcttcttcttcttcttcttctt 2820
DB 2761 tattcaagatgatagtcctcaggaagcttcttcttcttcttcttcttcttcttcttctt 2820
QY 2821 tatttgaagaaacggtcttcttgaatttaactt 2853
DB 2821 tatttgaagaaacggtcttcttgaatttaactt 2853

RESULT 2

ID AAX26841 standard; cDNA; 674 BP.

AC AAX26841;

DT 22-JUN-1999 (first entry)

DE Microspore-specific plant promoter cDNA clone M3.21.

KW Microspore-specific plant promoter; Brassica napus cv. Brutor;

KW gametophytic male sterility; inducible fertility; ss.
XX Brassica napus.
OS
XX
PN FR2768745-A1.
XX
PD 26-MAR-1999.
XX
PF 23-SEP-1997; 97FR-0011812.
XX
PR 23-SEP-1997; 97FR-0011812.
XX
PA (INRG) INRA INST NAT, RECH AGRONOMIQUE.
PI Drouaud J, Fourgoux A, Guerche P, Pelletier G;
XX WPI; 1999-217496/19.
DR
XX
PT Microspore-specific plant promoter from Brassica napus - and vectors
PT for producing plants exhibiting gametophytic male sterility with
PT inducible fertility
XX
PS Example 1; Page 14; 32pp; French.
XX
CC The present sequence represents a clone of a microspore-specific plant
CC promoter from Brassica napus cv. Brutor. The promoter is used in a
CC method for producing plants exhibiting gametophytic male sterility
CC with inducible fertility.
XX
SQ Sequence 674 BP; 155 A; 181 C; 124 G; 214 T; 0 other;

Query Match 16.2%; Score 462; DB 20; Length 674;

Best Local Similarity 81.9%; Pred. No. 4.2e-54;

Matches 614; Conservative 0; Mismatches 35; Indels 101; Gaps 3;

QY 2105 tcttgatgatctcttcttcaagctgtgtacatcccaagttacagcaagactaga 2164
DB 1 tcttgatgatctcttcttcaagctgtgtacatcccaagttacagcaagactaga 60
QY 2165 gtcatcaactaaccagaagctcttctctatcgcggaacttccgtcttcaaccacga 2224
DB 61 gtcatcaactaaccagaagctcttctctatcgcggaacttccgtcttcaaccacga 120
QY 2225 acattgacggtctcgtggtcgcggaagccttccctcgagccacttcgacttaacttc 2284
DB 121 acattgacggtctcgtggtcgcggaagccttccctcgagccacttcgacttaacttc 180
QY 2285 gtccatctgcccacagaagatgtaacagatgttgaacagaagaagaggttaagatag 2344
DB 181 gtccatctgcccacagaagatgtaacagatgttgaacagaagaagaggttaagatag 240
QY 2345 tttaataatgatacgtctgagacttcttccacagaagaagcgcgtatgtgacggaatgtg 2404
DB 241 tttaataatgatacgtctgagacttcttccacagaagaagcgcgtatgtgacggaatgtg 300
QY 2405 cgcgcgcatcaagaagatgtaacagaagatgtggaagaagcgttcttgaatcttcaatga 2464
DB 301 cgcgcgcatcaagaagatgtaacagaagatgtggaagaagcgttcttgaatcttcaatga 360
QY 2465 cccctcttgaacggtatgtcaagctacatgtctccacggttgttgaactactaac 2524
DB 361 cccctcttgaacggtatgtcaagctacatgtctccacggttgttgaactactaac 420
QY 2525 tctctcttcaagcgtccttcaatgctccttcttcaacagagctccttcaatgctccttc 2584
DB 421 tctctcttcaagcgtccttcaatgctccttcttcaacagagctccttcaatgctccttc 429
QY 2585 acatgctccttcaacagagctccttcaatgctccttcaatgctccttcaatgctccttc 2644
DB 430 -----acatgctccttc 441
QY 2645 acatgctccttcaacagagccttcaacagagccttcaacagagccttcaatgctccttc 2704

Db 442 acgggctccttaccatgctcctcctcaccagctccttaccatgccccctcaccagctcctt 501
OY 2705 actgcccccttgcagagctcctcaccgctcagtaga-tttgctatttgatagaattat 2763
Db 502 acgccccctcaccagctcctcaccgctcagtagatttgctatttgtagaattat 561
OY 2764 tcaagattgtagtctcctcagtagatttgctatttgctatttgtagaatttgctatt 2823
Db 562 tcaagtgtagtctcctcagtagatttgctatttgctatttgtagaatttgctatt 620
OY 2824 tttagaagaacccgcttggatttactt 2853
Db 621 tttagaagaacccgcttggatttactt 650

RESULT 3
AA26840
ID AAX26840 standard: cDNA, 497 BP.
XX
AC AAX26840;
XX
DT 22-JUN-1999 (first entry)
XX
DE Microspore-specific plant promoter cDNA clone M3.
XX
KM Microspore-specific plant promoter; Brassica napus cv. Brutor;
KM gametophytic male sterility; inducible fertility; ss.
XX
OS Brassica napus.
XX
PN FR2768745-A1.
XX
PD 26-MAR-1999.
XX
PE 23-SEP-1997; 97FR-0011812.
XX
PR 23-SEP-1997; 97FR-0011812.
XX
PA (INRA) INRA INST NAT RECH AGRONOMIQUE.
XX
PI Drouaud J, Fourgoux A, Guerche P, Pellelier G;
XX
DR WPI, 1999-217496/19.
XX
PT Microspore-specific plant promoter, from Brassica napus - and vectors
PT for producing plants exhibiting gametophytic male sterility with
PT inducible fertility
XX
PS Example 1; Page 13-14; 32pp; French.
XX
CC The present sequence represents a clone of a microspore-specific plant
CC promoter from Brassica napus cv. Brutor. The promoter is used in a
CC method for producing plants exhibiting gametophytic male sterility
CC with inducible fertility.
XX
SQ Sequence 497 BP; 107 A; 130 C; 70 G; 190 T; 0 other.

Query Match 13.8%; Score 393.2; DB 20; Length 497;
Best Local Similarity 98.0%; Pred.No. 7.7e-45;
Matches 398; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 2448 tttagatcttcatgaccctccttgcagctcagctatgtaagctacatgctcaccgct 2507
Db 1 tttagatcttcatgaccctccttgcagctcagctatgtaagctacatgctcaccgct 60
OY 2508 gttagatcttcatgaccctccttgcagctcagctccttcatgctccttccacagct 2567
Db 61 gttagatcttcatgaccctccttgcagctcagctccttcatgctccttccacagct 120
OY 2568 ccttaccatgctccttaccatgctccttaccagctccttaccatgctccttaccatgct 2627

Db 121 ccttaccatgctccttaccatgctccttaccagctccttaccatgctccttaccatgct 180
OY 2628 ccttaccatgctccttaccatgctccttaccagctccttaccatgctccttaccagctc 2687
Db 181 ccttaccatgctccttaccatgctccttaccagctccttaccagctccttaccagctc 240
OY 2688 ccttaccatgctccttaccatgctccttaccagctccttaccagctccttaccagctc 2747
Db 241 ccttaccatgctccttaccatgctccttaccagctccttaccagctccttaccagctc 300
OY 2748 tattgataagaattatcagaatgtagtgcctcagtaggagtttatttttctgtttt 2807
Db 301 tattgataagaattatcagaatgtagtgcctcagtaggagtttatttttctgtttt 360
OY 2808 aaatttgcgttatttattgagaacacgctccttggatttactt 2853
Db 361 aaatttgcgttatttattgagaacacgctccttggatttactt 406

RESULT 4
AAC34997
ID AAC34997 standard: DNA; 730 BP.
XX
AC AAC34997;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 8641.
XX
KM Hybridisation assay; genetic mapping; gene expression control;
KM protein identification; signal transduction pathway;
KM metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PE 25-FEB-2000; 2000EP-0301439.
XX
PF 25-FEB-1999; 99US-0121825.
XX
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126284.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 16-APR-1999; 99US-0128714.
PR 19-APR-1999; 99US-0129845.
PR 21-APR-1999; 99US-0130077.
PR 23-APR-1999; 99US-0130449.
PR 28-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 07-MAY-1999; 99US-0132486.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.

PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 30-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142309.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145274.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145813.
PR 27-JUL-1999; 99US-0145818.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.

PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151330.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154479.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160747.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160788.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.

Query Match	Best Local Similarity	6.1%	Score 174.8	DB 20	Length 2853
Matches 310	Conservative	68.6%	Pred: No. 1.9e-15	Mismatches 114	Indels 28
					Gaps
QY	793	atatacaagltatatactgttttccaataaagaataaataaataaataatag	852		
Db	1230	AGTATTAAATTAATATATATGTTTTATATATAAAAAGGTAAAAAAATTAATTAATGATAG	1171		
QY	853	tagtcaaaaaaaaaa-----aattaattttaccagcgt-----canaaac	898		
Db	1170	TTGTTCACAAAAAAATATTTTTTAAAAAATACCTTTGAACGTGCTGATGAACACATAAT	1111		
QY	899	actaaacctaaacctaatatliaacgttttagtugtaaacacctaaaccttgataaac	958		
Db	1110	CCTAAGTCATTAATCCCTTAACCATTAACCTCTGGGTAACCTTAAACCTTTTGATTAATTC	1051		
QY	959	ttaaacattaaacattaaacacacaaacctaaacctaaacctaaaccttaagtgct	1018		
Db	1050	CTAAACTATTAATTAATCAAAAAACCTAAACATTTTAAACACTTAAGCGTTTAGCTTAGATT	991		
QY	1019	taaaagttagtgcttttggattatagtttaggattatccaaaggtttaaagtttaacc	1078		
Db	990	TAGGCTTAGTGTATTATATGTTAATGTTAAGATTTTACCAAGCTTTAGGGTTTACCT	931		
QY	1079	aagagttatagtttaggattataacttaagatttagtggttttactagaagcttcaa	1138		
Db	930	AAAAGTTTAAATTAATTTAGGGTTTAGGTTTTAG-----TGTTTTNTATACCTGGTAAAA	879		
QY	1139	gtattttttaaataatcttttttggtaagaactacttttatacttttttttacct	1198		
Db	878	ATATT-----AATTTTTTTTTTTTGTGACACACTATATTTTATTTATTTTATTTTATCCT	825		
QY	1199	ttatatcaaaacataataataact	1230		
Db	824	TTTATTGAAAAACATAATATTAAGCTTGATTAAT	793		

ABL32315/c	
ID	ABL32315 standard; DNA; 6641 bp;
XX	ABL32315;
AC	
XX	
DT	26-MAR-2002 (first entry)
XX	
DE	Human immune system associated gene SEQ ID NO: 288.
XX	
KW	Human; immune system disease; cytosine methylation; antiasthmatic;
KM	antiarteriosclerotic; antiangiemic; cytostatic; nootropic;
KW	neuroprotective; anti-HIV; anticoagulant; ophthalmological;
KM	antihypertensive; antiarthritic; antididiabetic; antisporadic;
KW	antileukematory; cancer; eye disease; arteriosclerosis; anaemia;
KM	acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW	neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
KM	gene; ds.
OS	
XX	Homo sapiens.
PN	MO200200928-A2.
XX	
PD	03-JAN-2002.
XX	
PF	02-JUL-2001; 2001WO-EP07537.
XX	
PR	30-JUN-2000; 2000DE-1032529.
XX	01-SEP-2000; 2000DE-1043826.
PA	(EPIG-) EPIGENOMICS AG.
XX	
PI	Olek A, Piepenbrock C, Berlin K;
XX	
DR	WPI; 2002-130909/17.
PT	Nucleic acid comprising fragment of chemically modified gene, useful
XX	for diagnosis and treatment of diseases associated with abnormal
PT	cytosine methylation
PS	
XX	Claim 1; SEQ ID NO 288; 32pp + Sequence Listing; German.
CC	The present invention provides a number of human immune system associated
CC	genes which are modified by the methylation of cytosines. The sequences
CC	can be used in the diagnosis and treatment of immune system disorders,
CC	including eye diseases such as retinopathy, neovascular glaucoma and
CC	macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC	leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC	rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC	diseases. The present sequence is a gene of the invention.
XX	
SQ	Sequence 6641 BP; 2495 A; 22 C; 984 G; 3140 T; 0 other;
	Query Match 4.3%; Score 123.6; DB 24; Length 6641;
	Best Local Similarity 46.2%; Pred. No. 1.4e-08;
	Matches 604; Conservative 0; Mismatches 680; Indels 23; Gaps 5
QY	737 aaacaataaaattcattatgcatatcttatctttagagaagtgaacaaaataatta 796
DB	3626 AAACCTAAATTGAATCACTTACTTTCTTCTACTCGAANTTACTATTAATCTCA 3567
QY	797 tcaagttatatattgtttttccaataaagaataaaaaataataataaatagtagt 856
DB	3566 TTTTCTTTTCACAAAATTACACTCAATTTCTTAAATACTATCTATATATATTTGCC TT 3507
QY	857 tacaaaaaaaaaaaaataattatcttttacccaggctcamaaacactaaacctaaacctta 916
DB	3506 TCCTATATAAAAAAATTATATATACATTITA -CATATAATTTTTATATATAAATTCATATAA 3449
QY	917 aataataacatttaggiaaacaccataaccttgatataatcatcaacattaaacattaa 976
DB	3448 ACATTTAACCTCATAAAAAATTCAACAATATATATATATATATATATATATATATATATAT 3390

Oy	977	aacactaaaccctca-----atccctaaccctcaaacoccttaagtggtttaaatggttttagt	1032
Db	3388	CTCATTAATAATATTAATAATATAAATAATTAAACCTTATTCATATATTTAAAAACAAC	3329
Oy	1033	ttttgattataagttaggaattatcocaagglttaaggttatcccagaagtttagt	1092
Db	3328	CTATATATTTAAAATACCTTAATAAATAATAAATAAATAATATCTACTAATTTATTA	3269
Oy	1093	tagggattatgacctggagtttagtggttttctacgcagcgctccaagfatllttaaaa	1152
Db	3368	AAAATCCAAATATCTACTATTTATTAATAAACCTTAATAAAAAAAAACTTAATAATAATAA	3209
Oy	1153	atatcttttttgttcaca-a-actactattcttatatttttctactttatattaasa	1211
Db	3208	TATATTAAT	3149
Oy	1212	cataataataattaaactccalcgttttcataatgaaglygatgtlgaacattatl	1271
Db	3148	TAAATAAATTAATAATTTAA	3089
Oy	1272	tgttacaaaaaaatlgloacttttagagatcccatgcaaatattatttlttcaagcaa	1330
Db	3088	TATTA	3029
Oy	1332	aattaatgcaaaagtgcatatgactctcataataaattttatctatccaaatgcatlat	1391
Db	3028	TAAATAAATTAATAATTAACATTAATAATAATAATAATAATAATAATAATAATAATA	2969
Oy	1392	gtcgaacatgtgtcaattaaagaaccttaatatattcattatlcttttcttaacgt	1451
Db	2968	TTTTATATACAAAAAATATCAATAAAAAATATTAACATTTATTAATAATTAACCTTTAT	2909
Oy	1452	gtlaaaaatgccaagtaaataattatttagaagaacatgtagtaacatttggttcac	1511
Db	2908	TTTATTAATCTCCATATATCTT-----TCAATATATATATTAATTCATTTT	2864
Oy	1512	ttlaaagaatatgcaatatgaataaataacacaatttatgtatgacgaaccaaaatcat	1571
Db	2863	AACATATTAACCTCAAAATATACCTTAATAAATAATAATCAATTAATCTTCAACCAAAATTTAT	2804
Oy	1572	ccttaagaagtggaacgaagaataagtcgaogtlltgggaagactaactagggccaa	1631
Db	2803	TCAATATCCTAACATATTTCTTTAAAAAATATATTTTATATAAAAACATCCTTATATT	2744
Oy	1632	gtcalcaaaaactctctctgtattatcacaatccctacaactttagttagattataga	1691
Db	2743	TAAATTAAGTAATAATTAATCTTTATTAATAATTACAAAAAATAAATAATCTAACTTT	2684
Oy	1692	ccaacacatgatcatcatcatatagaatctctaaaaaactgaactagcgaalataa	1751
Db	2683	CCACATTAACACCTTAATAATCAATATCTTTATTAATAAATAAATAAATTAATTAATCTTTAA	2624
Oy	1752	atcttcttttatitacaacaaactcctataaanaacttatitatiaactaaacact	1811
Db	2623	TA	2564
Oy	1812	taattaaagaanaataaggacatgcatatcaataaataatatatgttattcttgaagtag	1871
Db	2563	TA	2504
Oy	1872	tgataaatatataabacagcccatatattatcaaaaataaataatataatttcogtagt	1931
Db	2503	TA	2444
Oy	1932	ccgaataacttactataaaatlcataaacaacatgtagatgat--tttattatata	1990
Db	2443	TACGTA	2384
Oy	1991	tatatataaaccotaaagccttacacatcgataaccctcoaaact	2037
Db	2383	TATATATATATCTCTAATATAAATACTAAATTAATACCAAAAAATATAT	2337

	RESULT	8	
ID	ABLJ34155/C		
ID	ABLJ34155 standard; DNA:	15548 BP.	
XX			
AC	ABLJ34155;		
XX			
DT	26-MAR-2002	(first entry)	
XX			
DE	Human immune system associated gene SEQ ID NO: 2128.		
XX			
KW	Human; immune system disease; cytosine methylation; antiasthmatic; antiartherosclerotic; antianemic; cytostatic; nootropic; neuroprotective; anti-HIV; anticoagulant; ophthalmological; antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia; acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy; neuromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene; ds.		
RN	Homo sapiens.		
OS			
PB	WO20020928-A2.		
PN			
PD			
PP	02-JUL-2001; 2001WO-EPO7537.		
PR	30-JUN-2000; 2000OE-1032529. 01-SEP-2000; 2000DE-1043826.		
PA	(EPIC-) EPIGENOMICS AG.		
PI	Olek A, Piepenbrock C, Berlin K,		
DR	WPI: 2002-130909/17.		
PT	Nucleic acid comprising fragment of chemically modified gene, useful for diagnosis and treatment of diseases associated with abnormal cytosine methylation		
PS	Claim 1; SEQ ID NO 2128; 32pp + Sequence Listing; German.		
CC	The present invention provides a number of human immune system associated genes which are modified by the methylation of cytosines. The sequences can be used in the diagnosis and treatment of immune system disorders, including eye diseases such as retinopathy, neovascular glaucoma and macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid leukemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis, CC rheumatoid arthritis, psoriasis and inflammatory/intestinal bowel diseases. The present sequence is a gene of the invention.		
SQ	Sequence 15548 BP; 4209 A; 247 C; 2903 G; 8189 T; 0 other;		
Query Match	: 4.3%; Score 122.4; DB 24; Length 15548;		
Best Local Similarity	44.7%: Pred. No. 2e-08;		
Matches	690; Conservative 0; Mismatches 833; Indels 19; Gaps 5		
DY	467 aattatacttgaggagaatatcaacttaacatgccaalatalagalagaanaaatattacc 526 		
DB	13157 AAAATATTATTTAATAAAATTAATTCCTAAAATATATATTAATAAAATTAATAAAAA 13098		
DY	527 ttgatgt-taacgtagtcggcgcagtlagagaagtccitgaataatatltagaaactcgctt 585 		
DB	13097 TAATAATATTAATAAATATTAATAAANTATATTAATTAATAAAATATATATATTACTAACATATACA 13038		
DY	586 tttcccacaagaatlaaalattlglatagtcaactlaaglttaacacatlgaaaataaaaaa 645 		
DB	13037 TATTATAAAATTAATAATCAATTAATCAGATTAATAAANAATCATATTAATACAAATTAATAATA 12978		
DY	646 aaattaaccaaaaatgagaaaacatgcatgcatcacacctcaacgcttttgaactatcc 705 		
DB	12977 AAATTATATTAATAAATAATATATTAATAAATATATTAATAATATATTAATGCGAAATATATAT 12918		

[illegible]

OY	1769	caaaatccctataaaaacatttatcatatcactaaaccatttaatttaaagaacaata	1828
Dd	11838	AAATAAAAATTAAAAATTAAAAATTAAAAATTAAAAATTAAAAATTAAAAA	117799
OY	1829	gggagccatcgatcacataaacaatatgttatttcctaagatagtgcatatcataat	1888
Dd	11778	TATTAATAATTAATAATTAATAATAATAATAATAATAATAATAATAATAATAA	11719
OY	1889	accgcgccatattttatcaaaataaaicaatatttttcgaagtcgatacatctact	1948
Dd	11718	TAAAATTAATTAATTAATAATTAATAATAATAATAATAATAATAATAATAAT	11659
OY	1949	aaatlcatbaaacacacatgtagatgatcatatttataatta 1990	
Dd	11658	AAAAATATAAAAAATTAAAAATTAAAAATTAAAAATTAAAA 11617	

RESULT	9
ABL32517/C	
ID	ABL32517 standard; DNA; 5930 BP.
XX	
AC	ABL32517;
XX	
DT	26-MAR-2002 (first entry).
XX	
DE	Human immune system associated gene SEQ ID NO: 490.
XX	
KW	Human; immune system disease; cytosine methylation; antiasthmatic;
KW	antiartherosclerosis; antianaemic; cytosolic; noctropic;
KW	neuroprotective; anti-HIV; anticomvulsant; ophthalmological;
KW	antirheumatic; antiarthritis; antidiabetic; antipsoriatic;
KW	antinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW	acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW	neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
gene;	ds.
XX	
OS	Homo sapiens.
XX	
PN	MO200200928-A2.
XX	
PD	03-JAN-2002.
XX	
PE	02-JUL-2001; 2001WO-EP07537.
XX	
PR	30-JUN-2000; 2000DE-1032529.
PR	01-SEP-2000; 2000DE-1043826.
XX	
PA	(EPIG-) EPIGENOMICS AG.
XX	
PI	Olek A, Piepenbrock C, Berlin K;
DR	WPI; 2002-130909/17.
XX	
PT	Nucleic acid comprising fragment of chemically modified gene, useful
PT	for diagnosis and treatment of diseases associated with abnormal
PT	cytosine methylation
XX	
PS	Claim 1; SEQ ID NO 490; 32pp + Sequence Listing; German.
XX	

The present invention provides a number of human immune system associated genes which are modified by the methylation of cytosines. The sequences can be used in the diagnosis and treatment of immune system disorders, including eye diseases such as retinopathy, neovascular glaucoma and macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis, rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel diseases. The present sequence is a gene of the invention.

Sequence	5930 BP; 2122 A; 24 C; 912 G; 2872 T; 0 other;
Query Match	4.2%; Score 121; DB 24; Length 5930;
Best local Similarity	45.8%; Pred. No. 3.2e-08;


```

OY 1062 aggttaaggttaccagaagtttaagggttaggattagcttaggtt 1121
DB 2972 TAAATAAATTAATTTCTCCCAACATCAAGTAACATTAAGTATCTCTT 2913
OY 1122 taactgaagcgttcaagaatttttaaaaaatatttttttgaactactattt 1181
DB 2912 CACTTAACACCTTCACATATACCAATACCTCTCTTAAACCTTAA 2853
OY 1182 attattttttacccttta-tattaacataataataatttaactccactggtt 1240
DB 2852 AACCATTTTAATAAATTTCTACCTTTTAACAAAATTTAAACCATTAACCTCTTA 2793
OY 1241 catattgaagtgcatgttaactattttttgttacaataaatgtcacttagaatt 1300
DB 2792 TACAATATCATGACCATTAATTAATAAATAATTAATCAAAACCATTAACAA 2733
OY 1301 ccaatgcaaaatttatttttttcaagctaaataatgtcgaagtgtacttata 1360
DB 2732 AAAACCTCACCTCTAATTAATAAATAATTAATAAATAAATAAATAAATAA 2673
OY 1361 aataattttattatctcaatgctatattgtcacaacatgtttaaattagaactta 1420
DB 2672 ACTAATTTTCAATCAAAAAAATAACATTATATTAATTAATTAATTAATTAATTA 2613
OY 1421 attataattcattatttttttctta----atctgttaaaatgtcagaatattat 1476
DB 2612 AATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2553
OY 1477 tgaagaagaattgtagtaatttcttcaattttttaaagaatcgaatgaaata 1536
DB 2552 CTAAATTAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2493
OY 1537 cacaattttatgtatgaagaaacttaaatctccttaagaagtggaagcaagaata 1596
DB 2492 CATCAACAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2433
OY 1597 gtaacggttttggaagacttaacttgcccaagtcacaacatcttctgttatta 1656
DB 2432 TCCTACCTACTATATATGAAATCTACCTTATCTTAATAAATAAATAAATAA 2273
OY 1657 tcaaaacttacaacttagttaggttaagaccacaacatgattatcatcatt 1716
DB 2372 ACATTAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2313
OY 1717 agaatttttaaaaaatctagcgaataaataatcttcttattatcaaaatcc 1776
DB 2312 TTTAAATCTTTTAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2253
OY 1777 ttataaaaacttattatataatacaaaactttaataaagaagaagagcat 1836
DB 2252 AATTCAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2193
OY 1837 ggaatacaataataatatttcttgaagatgtgataataataataccagtc 1896
DB 2192 AAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2137
OY 1897 atatttttcaaaaataataatttttgcagtcgaatacatcattactataatcat 1956
DB 2136 AATTAATTTCTTATCTTAATAAATAAATAAATAAATAAATAAATAAATAA 2077
OY 1957 aaaaacacatgtagatgtattttattatataataataaacccttaagccttac 2016
DB 2076 AATACCCGATCTTAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2017
OY 2017 caactgaatacaataaacttctctcgttgcgttaactcaaggtcgaagaata 2076
DB 2016 ATTAACCTTAATCAAAAACTTTTAAATATCTATTTCTAACCCCTCTAATTAATTTAA 1957
OY 2077 aaaaaacaatgaa 2089
DB 1956 AAAAAAATAAATAA 1944

```

```

RESULT 12
ABL33307/c
ID ABL33307 standard; DNA; 6175 BP.
XX
AC ABL33307;
XX
DT 26-MAR-2002 (first entry)
XX
DE Human immune system associated gene SRO ID NO: 1280.
XX
KW Human; immune system disease; cytosine methylation; antiasthmatic;
KW antiarteriosclerotic; antianaemic; cytosolic; nootropic;
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200200928-A2.
XX
PD 03-JAN-2002.
XX
PF 02-JUL-2001; 2001MO-EP07537.
XX
PR 30-JUN-2000; 2000DE-1032529.
XX
PR 01-SEP-2000; 2000DE-1043826.
XX
PA (EPIC-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2002-130909/17.
XX
PT Nucleic acid comprising fragment of chemically modified gene, useful
PT for diagnosis and treatment of diseases associated with abnormal
PT cytosine methylation.
XX
PS Claim 1: SEQ ID NO 1280; 32pp + Sequence Listing; German.
XX
CC The present invention provides a number of human immune system associated
CC genes which are modified by the methylation of cytosines. The sequences
CC can be used in the diagnosis and treatment of immune system disorders,
CC including eye diseases such as retinopathy, neovascular glaucoma and
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC diseases. The present sequence is a gene of the invention.
XX
SQ Sequence 6175 BP; 1506 A; 154 C; 1357 G; 3158 T; 0 other;

```

Query Match 4.2%; Score 119; DB 24; Length 6175;
Best Local Similarity: 45.9%; Pred. No. 5.9e-08;
Matches 594; Conservative 0; Mismatches 686; Indels 14; Gaps 5;

```

OY 816 tcaataaaaaagaataaaaaataaataaataaataaataaataa 875
DB 6083 TCTCAAAAAACAAAAAATAAATAAATAAATAAATAAATAAATAAATAA 6024
OY 876 tattttaccaggttcanaaaacacataaaccttaaatatataacttttgta 935
DB 6023 CATTAATAATTAACCTATTAATAAATAAATAAATAAATAAATAAATAA 5964
OY 936 aaccctaaaccttgataaatcttaaacataaacataaaccttaaac 995
DB 5963 ACAAAATAAATAATCTAATAATATTAATAAACAACGTAAATAAATAAATAATAC 5904
OY 996 taaactcaaaccttaagtttaaatgttttagtttttagtttagtttaggatt 1055

```


Oy 456 aaatgaagaaatttatatttgagaaataataattatgltcaatataagagaa 515
 Db 3267 ATATTTAAAAAAATTTTAAATTAATTAACATTTATTTAAACAAACAAATTAATATCA 3208
 Oy 516 aatatatncccttgatgctgcatgagtcgagtagaagatcttggaataatttga 575
 Db 3207 TCCACTACAAATTTAAAAACATACAAATTAATAATTTTAAACACTTAATTTCTTACA 3148
 Oy 576 gaacttgcccttctcgaagaataatgtaactgatactgaactgaactgaacatgaa 635
 Db 3147 ATTAACAAAAAAATTTAAACAAACAACTTTTAAATATATATACAAATTTTAAAAAAA 3088
 Oy 636 attaaaaaaataatacaaatagaaaaactga--tagagactcccttaagct 692
 Db 3087 AAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3028
 Oy 693 ttggaacttctcttggttcaaccccttaacccctgaagtcacacacataaattca 752
 Db 3027 TTTAAACGTCAAAAATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2968
 Oy 753 ttatgcatatctcatatctttagaagtggaacaaaataatataatgatatatg 812
 Db 2967 ATATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2908
 Oy 813 ttccaataaagaataaaaaataaataataatagtagtatacaaaaaaaaat 872
 Db 2907 ATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2848
 Oy 873 taatatcttaccagcgctcaaaaaaacctaaacctaacttaactttag 932
 Db 2847 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2794
 Oy 933 gtaaacctaacccttgatataatcttaaacatlaaacatlaaacctaaccctaa 992
 Db 2795 -TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2737
 Oy 993 tccaaacctaaacccctgaagtggttaagtggttagtggttagttagttagttagt 1052
 Db 2736 ACCCAACGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2677
 Oy 1053 ttatccaaaggttaaggttaccgaaggttagtggttaggaatgaacttagat 1112
 Db 2676 AAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2619
 Oy 1113 ttagtggttctgagcgctgaacatatttcttgtaaaaaaatatttcttgtaaac 1172
 Db 2618 AAACCTCTTACTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2561
 Oy 1173 actatcttattatcttcttacccttcttaataaaaaaacatataacttactc 1232
 Db 2560 TCTTACTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2501
 Oy 1233 atcgttctcatatgaatgcatctgtaacatatttcttgtaaaaaaatgctac 1292
 Db 2500 ATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2441
 Oy 1293 ttagaatccaaatgaatattatatttctgaagc-----aaatgaatgaagtc 1348
 Db 2440 CTTAAACCTACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2381
 Oy 1349 atgactctataataattatattatctcaaatgcatatctgtaacaaatgtaact 1408
 Db 2380 AAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2321
 Oy 1409 aatgaactaatatattatcttatttcttcttaactctgtaaaaatgtaagta 1468
 Db 2320 AAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2261
 Oy 1469 aaatttattagaagaactgagtaaatattggttcttctttaaagaatagata 1528
 Db 2260 AAATCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2207
 Oy 1529 tgaataacaaatttatgtatgtaaacctlaaaaatctcctaagaagtgtaagc 1588

Db 2206 TTTAAACAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2147
 Oy 1589 aagaatgaactgaacttcttggaagctgaacttgcccaagctcaaatcttct 1648
 Db 2146 AACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2091
 Oy 1649 tgaattctcaaaatctcttaacaaatagtagtaagtaataagcaacacatgatac 1708
 Db 2090 CCTATTATCTACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2031
 Oy 1709 atcatatgatatctcaaaaatctagcgaaatgaatlaaaccttcttcttattat 1768
 Db 2030 ATAAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1971
 Oy 1769 caaaatcctataaaaactattatataatacaaaaacttaataaagaataa 1828
 Db 1970 AAACAAACAACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1911
 Oy 1829 gggacacatgatacaaaaataatagtagtattcttcaagatgataatataat 1888
 Db 1910 TAAATATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1851
 Oy 1889 accagtcataatatctcaaaaataaataatcttctgtagtcga-----taacatata 1944
 Db 1850 TTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1791
 Oy 1945 ctataatctataaaacacatgtagatgataattatattatataataataacc 2004
 Db 1790 TCTTTATTTAAAAAATTTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 1731
 Oy 2005 ctacgcctcaccacacgcgaataacataaaccttctctcggttgcgaacga 2064
 Db 1730 AAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1671
 Oy 2065 ttggaagaatgaataa 2079
 Db 1670 ATCTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1656

RESULT 15
 ABL34155
 ID ABL34155 standard; DNA; 15548 BP;
 XX
 AC ABL34155;
 XX
 AC
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Human immune system associated gene SEQ ID NO: 2128.
 XX
 KW Human; immune system disease; cytosine methylation; antiasthmatic;
 KW antiarteriosclerosis; antianemic; cytosolic; noctropic;
 KW neutrophilic; anti-HIV; anticonvulsant; ophthalmological;
 KW antineuritic; antiarthritic; antidiabetic; antipruritic;
 KW antinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
 KW acute myeloid leukemia; Alzheimer's disease; AIDS; epilepsy;
 KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
 KW gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN W0200200928-A2.
 XX
 PD 03-JAN-2002.
 XX
 PE 02-JUL-2001; 2001MO-EP07537.
 XX
 PR 30-JUN-2000; 2000DE-1032529.
 XX
 PR 01-SEP-2000; 2000DE-1043826.
 XX
 PA (EPIC-). EPIDEMIOLOGY AG.
 XX
 PI Olek A, Piepenbrock C, Berlin K;

XX WPI; 2002-130909/17.

PT Nucleic acid comprising fragment of chemically modified gene, useful
PT for diagnosis and treatment of diseases associated with abnormal
PT cytosine methylation

PS Claim 1; SEQ ID NO 2128; 32pp + Sequence Listing; German
xx

CC The present invention provides a number of human immune system associated
CC genes which are modified by the methylation of cytosines. The sequences
CC can be used in the diagnosis and treatment of immune system disorders,
CC including eye diseases such as retinopathy, neovascular glaucoma and
CC macular degeneration, arteriosclerosis, anemia, cancer, acute myeloid
CC leukemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC rheumatoid arthritis, psoriasis and inflammatory/allergic bowel
CC diseases. The present sequence is a gene of the invention.

Sequence 15548 BP; 4209 A; 247 C; 2903 G; 8189 T; 0 other;

Query Match	4.28	Score 118.6	DB 24	Length 15548
Best Local Similarity	44.28	Pred. No. 6.4e-08		
Matches 748	Conservative 0	Mismatches 928	Indels 18	Gaps 6

OY	303	tataacgaggtttgaagttatcctttttaaacttaataataacagagctttac	362
OY	11616	tttaatttttaatttttaattttatatttttaatttttaatttttaattttttt	11675
OY	363	ttatbaatttggncatgtatcatcaggttttagcncctttttttgancgcgtgatt	422
OY	11676	ataatttatcttattttattttttattttattttattttatttttttttttttt	11735
OY	423	ggtttatgcttatttgaatgngcnaagtaagaagaagaacatttatatttggaga	482
OY	11736	tatatatttatcatatattttatttttaattttttatattttatttttttttttt	11795
OY	483	aaataaatttaataatgttcaaatatagagagaataatnccctgaatgttaccgtatg	542
OY	11796	atttttatatt	11855
OY	543	gatcgagttagaagatcttttgaataataatttggagaacttgcttttccaaaagtaaa	602
OY	11856	tttatttttatattttatattttttttttttttttttttttttttttttttttttta	11915
OY	603	tattgataatgaacttaagttaacacagtaaacataaaaaaatttaacnaaataga	662
OY	11916	ttttatttttatatt	11975
OY	663	aaaaacgtatgtgacctaacccttcaacggttttgaacttattcttgyttcacccctaaa	722
OY	11976	ataatgtatataattttatt	12033
OY	723	cctcgaatgtccacaaacaataaatttccatttgcattgtcatctatcttttttaggaagt	782
OY	12034	tttatatgttattttttatatgttatttttttttttttttttttttttttttttttttttt	12093
OY	783	gaacacaaatatctacaagtataatattatgttttttccaataaaagaataaaaaataaa	842
OY	12094	ttatttttatatgttatt	12153
OY	843	aaataatagtagttacaaaaaaataataatttttttccagcgctnaaaaaacata	902
OY	12154	atagttatataatataatgttataataatcgtttatataatataatagttataataatagttat	12213
OY	903	aaacctaaacctaaatatataaacttttagtgaacaccttaaaccttttgataaatcttaca	962
OY	12214	atatataatgtatataatataatagttatataatagttataataatagttataatagttat	12269
OY	963	aaattaaacattaaacacctaaaccttaaaccttaaaccttaaaccttaaaccttaaacct	1022
OY	12270	atatataatgtatataatataatagttatataatagttataatagttataatataatagttataat	12329

QY	1023	tgtttagtgcttttgatttatagttttaggatttaccgaaggtttaggtttaccgaaga	1082
Dd	12330	atagtctatatatatataatgcttatacatatagtttatatatatacagtttatataat	12389
QY	1083	gtcttaggttttagggattatgacttaggaatttagtggtttactagcagcgttcaagta	1142
Dd	12390	atgtatgtatataatgcttatatatatatatttttag---tatatatgtatatatatata	12446
QY	1143	tttttaaaaaaatctttttttgtgaacaactactattttattttttttttaccctttta	1202
Dd	12447	tttttagtatatacttttatatatatatatttttagtatacatattttatatataatctt	12506
QY	1203	tattaaaaacaataatatttaatttaaccocactgtttccatatttagtgcactgttaaca	1262
Dd	12507	atgtatatatttttatatatatatatttttagtatatattttttatatatatatttttagt	12566
QY	1263	ttattttttgttacaanaaaaaatgtccacttagaattcccaatgtcaaaattattatctt	1322
Dd	12567	atatatttttatatatataatatttttagtataatgttataatatatatatatatatat	12656
QY	1323	ttcagctaaaatttaattgtcaagtgcatgtgaacctataaataattttatttcccaat	1382
Dd	12627	ttcgtatatatatatttttatatatatatatatatttatcttgcgtatatatatatttta	12686
QY	1383	gctaatatgtgcacaacatgtgcaattgaattagaacct-aaattatatcttatttttt	1441
Dd	12687	tatatatatatatatatattttcgtatatatatatatattttatatatatatttatatacttc	12746
QY	1442	cttaactcgtgttaaaaatgtccaagtlaaaatttattttagnaaagcaattgagtaatttt	1501
Dd	12747	gatatatatatttttatatatatatatatatatatttcgtatatatatatttcgtlat	12806
QY	1502	gtttcattttttaaaagatatcgaatatgaataatacacaaattttattgtatagaacct	1561
Dd	12807	aatatatatatatatatatatatatttcgtatatatatatttcgtatatatatattata	12866
QY	1562	aaaaatctccctaagaagtgtaacgcaagaataagccaagttcgtgggaaagcnaact	1621
Dd	12867	tatatattcgtatatatatatatatttcgtatatatatatttatatatatatatttcgtatatatatatt	12926
QY	1622	atgcccacaagtcacacaacactcttcctgtatttatccaanaactctacaanaattgtag	1681
Dd	12927	tcgtatatatatatatatatatttttatatatatatat-----tttatatatatttta	12979
QY	1682	agttaatagaccaacacactgtatcatcatatagaatctccaaaanaattactagcg	1741
Dd	12980	tatatatatgttatatatatgtatttatatatatagtataatgttatatatattttatag	13039
QY	1742	ataaattaaaactcttcttttattatccaacactcttaaaaactcttatatatact	1801
Dd	13040	tattctgtatatatatatttttatatatatatatttttatatatatttttatatatatt	13099
QY	1802	aaaaacaatttttaataaagaanaataagggaccatggaagtaacttaaaaataatatgttatt	1861
Dd	13100	ttatatatatatttttatatatatatatatatttttagatatatatatttttatatatatatatttttag	13159
QY	1862	cttaagaatagtatatatataataataaccagtcacatatattctacaanaataaataat	1921
Dd	13160	atatatatattttatatatatatttttagatatatatatttttatatatatatatttttagatatatt	13219
QY	1922	tttcgttagtcggaataactaattactataaattcatanaaacacacatgtagaatgtatt	1981
Dd	13220	atttatatatatatat-acttttagatatatatatttttatatatatatatttttagatatatttt	13278
QY	1982	atttatatatatat- 1995	
Dd	13279	ataatatatatattc 13292	

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 28, 2002, 09:02:55 ; Search time 2155.85 Seconds
(without alignments)
17861.545 Million cell updates/sec

Title: US-09-509-188-3

Perfect score: 2853
Sequence: 1 ggaaccacaagaacacg.....ccgtcttgatttaactt 2853

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 13736207 segs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estlba:*
2: em_estlhum:*
3: em_estlin:*
4: em_estlmu:*
5: em_estlov:*
6: em_estlpl:*
7: em_estlro:*
8: em_hlc:*
9: gb_estl1:*
10: gb_estl2:*
11: gb_hlc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	688.2	24.1	758	12	BH540198
C 2	646.6	22.7	748	12	BH540204
C 3	445.8	15.6	783	12	BH460089
C 4	378.2	13.3	801	12	BH563990
C 5	229	8.0	872	12	BH478681
C 6	228.6	8.0	554	12	BH582777
C 7	228.2	8.0	739	12	BH467584
C 8	227	8.0	739	12	BH467584
C 9	226.4	7.9	614	12	BH581638
C 10	217.8	7.6	790	12	BH603601
C 11	206.8	7.2	836	12	BH570528
C 12	205.4	7.2	810	12	BH477804
C 13	205	7.2	810	12	BH441036
C 14	202.2	7.1	872	12	BH478681
C 15	199.6	7.0	430	12	BH544194
C 16	199.2	7.0	790	12	BH603601
C 17	192.4	6.7	836	12	BH570528

C 18	192.2	6.7	677	12	BH007803
C 19	192.2	6.7	677	12	BH007804
C 20	189.4	6.6	462	12	BH011048
C 21	189.2	6.6	813	12	BH250621
C 22	186.4	6.5	504	9	AV549076
C 23	184.2	6.5	626	12	BH558350
C 24	182	6.4	761	12	BH425956
C 25	180.2	6.3	462	12	BH011048
C 26	179	6.3	787	12	BH582306
C 27	177	6.2	604	12	BH564610
C 28	177	6.2	765	12	BH479185
C 29	176.6	6.2	605	12	BH560885
C 30	175.8	6.2	753	12	BH476224
C 31	172.6	6.0	703	12	BH606245
C 32	172.6	6.0	808	12	BH581818
C 33	172	6.0	812	12	BH455873
C 34	171.8	6.0	761	12	BH425956
C 35	171.6	6.0	727	12	BH570602
C 36	171.4	6.0	785	12	BH532651
C 37	167.4	5.9	748	12	BH540204
C 38	167	5.9	686	12	BH454053
C 39	166.8	5.8	796	12	BH606626
C 40	163.6	5.7	721	12	BH556962
C 41	163	5.7	614	12	BH581638
C 42	162.4	5.7	799	12	BH421537
C 43	161.8	5.7	746	12	BH592365
C 44	160.8	5.6	576	12	BH456908
C 45	160.8	5.6	603	12	BH010155

ALIGNMENTS

RESULT 1
BH540198/c
LOCUS BH540198 758 bp DNA linear GSS 14-DEC-2001
DEFINITION BCGT Brassica oleracea genomic clone BCGT58TR
ACCESSION BH540198
VERSION BH540198.1 GI:17785802
KEYWORDS GSS.
SOURCE Brassica oleracea.
ORGANISM Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
1 (bases 1 to 758)
REFERENCE
TOWN, C.D., Van Aken, S., Uterback, T. and Fraser, C.M.
Whole genome shotgun sequencing of Brassica oleracea
Unpublished (2001)
Other GSSs: BCGT58TR
COMMENT
Contact: Chris Town

TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: ctown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TF
Class: sheared ends.
Location/Qualifiers
1..758
/organism="Brassica oleracea"
/strain="TO1000DH3"
/db_xref="taxon:3712"
/clone_lib="BOGT"
/note="Vector: pHS01; site 1: BstXI; 2-3 kb sheared genomic DNA inserted into pHS01 using BstXI linkers"

BASE COUNT 228 a 106 c 155 g 269 t
ORIGIN

|||||
Db 671 CAAATTTTATTTTATTTTACCTTTTATTTTATTT-AAAACATATATATTTTAACTCCA 729
Oy 1234 tctgttcatatata 1248
|||||
Db 730 TCTGTTTCATATTTAA 744
RESULT 3
BH460089 783 bp DNA linear GSS 13-DEC-2001
LOCUS BOGSK21TF BOG Brassica oleracea genomic clone BOGSK21, DNA
DEFINITION sequence.
ACCESSION BH460089
VERSION BH460089.1 GI:17649834
KEYWORDS GSS.
SOURCE Brassica oleracea.
ORGANISM Brassica oleracea.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
1 (bases 1 to 783)
Town, C.D., Van Aken, S., Uterback, T. and Fraser, C.M.
Whole genome shotgun sequencing of Brassica oleracea
Unpublished (2001)
Other_GSSs: BOGSK21TF
COMMENT Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TF
Class: sheared ends.
FEATURES
Location/Qualifiers
source 1..783
/organism="Brassica oleracea"
/strain="TO1000DH3"
/db_xref="taxon:3712"
/clone="BOGSK21"
/clone_1lb="BOGSK"
/note="Vector: PHOS1, Site.1: BstXI; 2-3 kb sheared
genomic DNA inserted into PHOS1 using BstXI linkers"
BASE COUNT 228 a 183 c 161 g 211 t
ORIGIN
Query Match 15.6%; Score 445.8; DB 12; Length 783;
Best Local Similarity 92.3%; Pred. No. 2,4e-38;
Matches 481; Conservative 0; Mismatches 37; Indels 3; Gaps 1;
Oy 2101 ttgtcttgctatgacccctcttaagctggtgacatcccaagttaagcaagaac 2160
|||||
Db 1 ttgtcttgctatgacccctcttaagctggtgacatcccaagttaagcaagaac 60
Oy 2161 tagagtcacaaactaaacgaagctcttctatcgcgcactactcgtttcacccca 2220
|||||
Db 61 TAGAGTCATCACTAACCAAGTCTTCTTATCGGGGCACTTGCCTTCACCCCA 120
Oy 2221 agcaacattgagcgttcgltggtctccggaagaagccttcctgagggcaccttcgactaa 2280
|||||
Db 121 AGCCACATATGGAGTTCCTGCGTCCGGAAGACCTTCCCTGAGGCCACTTCGACAA 180
Oy 2281 ctccgttcacatctgcacagaagatgacagatgcttgaaacgaagaagggtgagta 2340
|||||
Db 181 CTCGCTTCATCTGCGCACAGGAAGTCAACAGATGCTTGTCGCCAAGAAGAGGTAGGTA 240
Oy 2341 catgttttaatatcgatcgatgaacttctcacccagaagaagcgttatgtgagtaat 2400
|||||
Db 241 CATGTTTATATATATCGTTGAGACTTTCTCTCCACAGGAAGCGCGCTATTGATGATCGAAT 300
Oy 2401 gtctgcgcgcatcaagaagatgtaacaaagattgtgagaagaccgtcttgatcttcc 2460

|||||
Db 301 GTTGCCGCCCGCATCAAGATGACAAAGATTGTGAGAGACCGCTTTGGATCTTTC 360
Oy 2461 atgaccctcttgaccgctatgtaagctatcatgctccaccgttgttgatctact 2520
|||||
Db 361 ATGACCCCTTCTTGACAGGCTATGTCAACTATGCTCCACCGGTGTGGATCTACTT 420
Oy 2521 caactcctcttcacagagctccttaccatgctcctcttcacagagctcctacatgctc 2580
|||||
Db 421 CACCTCTCTCTTACATGCTCTTCACAGGCTC---CTTACATGCTCTTCACAGGCTC 477
Oy 2581 ctacacatgctccttcacagagctcctttaaagctcctta 2621
|||||
Db 478 CTTTACTGCCCGCTTCACAGCTCCTCCTCACCGGCTCAGTGA 518
RESULT 4
BH563990 801 bp DNA linear GSS 14-DEC-2001
LOCUS BOGVZ70TR BOGV Brassica oleracea genomic clone BOGVZ70, DNA
DEFINITION sequence.
ACCESSION BH563990
VERSION BH563990.1 GI:17815830
KEYWORDS GSS.
SOURCE Brassica oleracea.
ORGANISM Brassica oleracea.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
1 (bases 1 to 801)
Town, C.D., Van Aken, S., Uterback, T. and Fraser, C.M.
Whole genome shotgun sequencing of Brassica oleracea
Unpublished (2001)
Other_GSSs: BOGVZ70TR
COMMENT Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TR
Class: sheared ends.
FEATURES
Location/Qualifiers
source 1..801
/organism="Brassica oleracea"
/strain="TO1000DH3"
/db_xref="taxon:3712"
/clone="BOGVZ70"
/clone_1lb="BOGV"
/note="Vector: PHOS1, Site.1: BstXI; 2-3 kb sheared
genomic DNA inserted into PHOS1 using BstXI linkers"
BASE COUNT 231 a 176 c 177 g 217 t
ORIGIN
Query Match 13.3%; Score 378.2; DB 12; Length 801;
Best Local Similarity 99.2%; Pred. No. 3e-31;
Matches 380; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Oy 2114 galcctcttcttaagctggtgacatcccaagttaagcaagaagatgacatcaac 2173
|||||
Db 1 GATCTATTCTTTAAGCTGTGTCATCTCACTCAAGATTACAGCAAGAACTAGATCATCAAC 60
Oy 2174 taaccaagagctcttctatcgcgcactactcgtttcaccccaagaagaattggcc 2233
|||||
Db 61 TAACCAAGAGCTCTTCTTATCGGGCACTTACCTGCTTTCACCCCAAGCAACATTTGGCC 120
Oy 2234 gtctcgtgctcgcggaagaagccttcctgagggcaccttcgactaactcgtttcatct 2293
|||||
Db 121 GTTCGCTGCTCGGGAAGACCTTCCTCGAGGTCACTTCGCACTAATCTCCGTTCCATCT 180
Oy 2294 gccacagaagatcaccaagatgcttgaaacgaagaagagtagtatacgttttaatga 2353

Db 181 GCCACAGAGATGACCAATGCTTGAAGACAGACAGAGTACGTTATTAATGA 240
OY 2354 tatcgtaagactttctcaccagaagccgctattgagtcggaagtgccgcgat 2413
Db 241 TATGCTTACAGCTTCTTCCACCAAGAAAGCCGCTATTGATCGAATGTTGCCCGGAT 300
OY 2414 caagaagatgacaaagattgtgagaagccgtcttggatcttccatgacccttct 2473
Db 301 CAAGACAGATGACAAAGATGTGTAGAGACCGCTTTGATCTTTCATGACCCCTTCTT 360
OY 2474 gaccggtatgtcaagctacatt 2496
Db 361 GACCGCTATGTCAAGCTACATT 383
RESULT 5
BH478681 872 bp DNA linear GSS 13-DEC-2001
LOCUS BH478681
DEFINITION HOGT236TF BOGT Brassica oleracea genomic clone BOGT236, DNA
sequence.
ACCESSION BH478681
VERSION BH478681.1 GI:17686785
KEYWORDS GSS.
SOURCE Brassica oleracea
ORGANISM Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE 1 (bases 1 to 872)
AUTHORS Town,C.D., Van Aken,S., Utterback,T. and Fraser,C.M.
TITLE Whole genome shotgun sequencing of Brassica oleracea
JOURNAL Unpublished (2001)
COMMENT Other GSSs: BOGT236TR
Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TR
Class: sheared ends.
FEATURES
location/Qualifiers
source 1..872
/organism="Brassica oleracea"
/strain="TO1000DH3"
/db_xref="taxon:3712"
/clone="BOGT236"
/note="Vector: PHOS1; Site 1: BstXI; 2-3 kb sheared
genomic DNA inserted into PHOS1 using BstXI linkers"
BASE COUNT 313 a 135 c 109 g 315 t
ORIGIN
Query Match 8.0%; Score 229; DB 12; Length 872;
Best Local Similarity 72.2%; Pred. No. 1.3e-15;
Matches 390; Conservative 0; Mismatches 122; Indels 29; Gaps 6;
OY 735 ccaacataaattcattatgcatatcattatcatttgaagaagtgaaacaaat 794
Db 4 CCAACCAATGATTCATCTACTCTATATTCGATATTTTGTAGAAAGCTTACAAATAT 63
OY 795 tatcaagtatatattgttttcaataaagaataaanaataaataaataataata 854
Db 64 TGTCAAGTATATATTACGTTTAAATAAAGATTAATAAATAAATAAATAAATAA 123
OY 855 gttaacaaaaaataaataaataaataaataaataaataaataaataaataaata 910
Db 124 GTTACAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 183
OY 911 accctaataaataaataaataaataaataaataaataaataaataaataaata 965

Db 184 CTCCTGATATACATCTTAACCTTTGGTAAATCTTAACCTTTGATTAATCTTAAC 243
OY 966 ttaacatlaaacaacaaacaaacaaacaaacaaacaaacaaacaaacaaacaaac 1025
Db 244 TAAATCA--AAACACATAAATCTTAACCCCAATCTTGTAGATTTTGTGTTAATGTT 301
OY 1026 ttagtctt-----ttgattatagtttagattatccaaagtttaagttt 1074
Db 302 TTGATTTAGAGTTTACGATTTTATCCAAAGGTTAGGCTTACCTTAAGGTTAGGCTT 361
OY 1075 acccaagttttagtttaggtatgacttagattagtttagtttaggttagt 1134
Db 362 ATCCAGGTTTAAAGATTTGATTTAGAGTTTAAAGGTTTGTGTTTCTCAGACGTT 421
OY 1135 caaagtatttttaaaaaa-----tattttttgtaacactactatatttatt 1188
Db 422 AAAAATAGCTTTTAAATAAATTTCTAATTTTGTGTTTACATTTACTATTTTATTTT 481
OY 1189 ttttaacttttatataaacaataataataataataataataataataataata 1248
Db 482 TTTTAA-TTTTGTGTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATA 540
RESULT 6
BH582777 554 bp DNA linear GSS 15-DEC-2001
LOCUS BOHNA49TR BOHA Brassica oleracea genomic clone BOHNA49, DNA
DEFINITION
sequence.
ACCESSION BH582777
VERSION BH582777.1 GI:17835234
KEYWORDS GSS.
SOURCE Brassica oleracea
ORGANISM Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE 1 (bases 1 to 554)
AUTHORS Town,C.D., Van Aken,S., Utterback,T. and Fraser,C.M.
TITLE Whole genome shotgun sequencing of Brassica oleracea
JOURNAL Unpublished (2001)
COMMENT Other GSSs: BOHNA49TF
Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TR
Class: sheared ends.
FEATURES
location/Qualifiers
source 1..554
/organism="Brassica oleracea"
/strain="TO1000DH3"
/db_xref="taxon:3712"
/clone="BOHNA49"
/note="Vector: PHOS1; Site 1: BstXI; 2-3 kb sheared
genomic DNA inserted into PHOS1 using BstXI linkers"
BASE COUNT 204 a 88 c 70 g 192 t
ORIGIN
Query Match 8.0%; Score 228.6; DB 12; Length 554;
Best Local Similarity 73.9%; Pred. No. 1.8e-15;
Matches 343; Conservative 0; Mismatches 90; Indels 31; Gaps 3;
OY 718 cttaaccttaagtttaccacaaataaatttcattatgcatatcattatcattttag 777
Db 69 CTTCACCTCTAGTTCACCAACCAATAGATTTTCCTTATTTCAAAATTCGGATATTTTGG 128
OY 778 aaagtgaaacaaataataataaataaataaataaataaataaataaataaata 837

129 AACCAAAACAAATATCTCAAGTATATATGTTTAAATAAAGTAAAAA 188
838 aataaanaaataagtagtacaataaataaataatattttccagcgctcnaa--- 894
189 AATAGTAGTACAAAATAAATAATTTTAAATTTTAAAGTCTGCTGCAAAACACTA 248
895 -----aaacataaaccataaaccataataaacttttagtaaacct 941
249 ATCCCTAAACCTTAATCGGAACCTTAATCGGAACCTTAATCGGTAACCTCT 308
942 aaaccttggataaactc-----aaacataaaccataaaccataaacc 987
309 AAACCTTGATTAATCAATACTAACTAACTAACTAACTAACTAACTAACTAACT 368
988 ctaaaccttaaaccttaaaccttaaaccttaaaccttaaaccttaaaccttaaacct 1047
369 CTAAACCTTAACCTTAACCTTAACCTTAACCTTAACCTTAACCTTAACCTTAACCT 428
1048 taggattatccaaaggttaaggtttaccgaaggtttataggttttaggattagctt 1107
429 TAGATTATTAACAAAGCTTTAGGCTTTACCAAAAGTTTA-GGTTTAAGGATTAGATT 487
1108 aggatttagttagttagttagttagttagttagttagttagttagttagttagttagt 1151
488 AGGGTTAGTGTCTTCTAAGCAAGCTTAACAAATATTTTGTGTA 531

RESULT 7
BH467584 739 bp DNA linear GSS 13-DEC-2001
LOCUS BOHRG31TF BOHR Brassica oleracea genomic clone BOHRG31, DNA
DEFINITION
sequence.
ACCESSION BH467584
VERSION BH467584.1 GI:17670716
KEYWORDS GSS.
SOURCE Brassica oleracea.
ORGANISM Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
1 (bases 1 to 739)
REFERENCE
TOWN,C.D., Van Aken,S., Uterback,T. and Fraser,C.M.
TITLE Whole genome shotgun sequencing of Brassica oleracea
JOURNAL Unpublished (2001)
COMMENT Other GSSs: BOHRG31TR
Contact: Chris Town

TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TF
Class: sheared ends.
Location/Qualifiers
1. 739
/organism="Brassica oleracea"
/strain="TO1000DH3"
/db_xref="taxon:3712"
/clone="BOHRG31"
/note="Vector: pHOS1, Site 1: BstXI; 2-3 kb sheared
genomic DNA inserted into pHOS1 using BstXI linkers"
BASE COUNT 263 a 95 c 99 g 282 t
ORIGIN

Query Match 8.0%; Score 228.2; DB 12; Length 739;
Best Local Similarity 68.8%; Pred. No. 1,7e-15;
Matches 407; Conservative 0; Mismatches 139; Indels 46; Gaps 5;
QY 719 taacctaagttcccaataaattcatattatgcatattcatattttaga 778

129 TGAACCTATGTTTACCAACCAATGAGATTGTTATTTCAATTTGATCTTTAAG 188
779 aagtgcaaaaaatattatcaagttatattatgttttcaataaagaataa---aa 894
189 AATAGTAGTACAAAATAAATAATTTTAAATTTTAAAGTCTGCTGCAAAACACTA 248
835 ataaataaataaataagtagtacaataaataa-----aaataatattttccag 887
249 ATTAATTAATAAATAAGTAGTACAAAATAAATGTTTTTAATAATTTTATGCGT 308
888 cgtca-naaaacataaaccataaaccataaataaacttttagtaaaccttaaac 946
309 CATCACCAAAACCTAATCTAATCCCTTAATCCCTTAATCCCTTAATCCCTTAATCC 367
947 ttggtaaatcttaaacattaacattaaacataaaccataaaccataaaccataaacc 1006
368 CTGGAATAAATCTTAATGTTTAATCAAAACACTTAACACTTAACACTTAACACTTA 427
1007 cccttaagtttaaaatgttttagttagttagttagttagttagttagttagttagttagt 1066
428 GGGTTAATGTTTATGATTATTTTATGATTATTTATGATTATTTATGATTATTTATG 487
1067 taagtttaccgaaggttagttagttagttagttagttagttagttagttagttagt 1126
488 TATGAT-----TTAGATTATTAATGTTTATTTG 514
1127 agagcttcaagattttttaaataatattttttagtaaacactatttttattt 1186
515 ACACGCTTAATAAATATTTTAAATTTTATTTATTTATTTATTTATTTATTTATTT 574
1187 ttttttacccttttlatataaataaataaataaataaataaataaataaataaataaataa 1246
575 AGTTTACTGCTTAATTTAAACATATATTAATTAATTTTGTGTTCCATTTT 634
1247 aagtgcaatgttaacattattttttagttagttagttagttagttagttagttagttagttagt 1298
635 AAAGATATGCAATATCAATTAATTAATTTTATTTATTTATTTATTTATTTATTTATTT 686

RESULT 8
BH467584 739 bp DNA linear GSS 13-DEC-2001
LOCUS BOHRG31TF BOHR Brassica oleracea genomic clone BOHRG31, DNA
DEFINITION
sequence.
ACCESSION BH467584
VERSION BH467584.1 GI:17670716
KEYWORDS GSS.
SOURCE Brassica oleracea.
ORGANISM Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
1 (bases 1 to 739)
REFERENCE
TOWN,C.D., Van Aken,S., Uterback,T. and Fraser,C.M.
TITLE Whole genome shotgun sequencing of Brassica oleracea
JOURNAL Unpublished (2001)
COMMENT Other GSSs: BOHRG31TR
Contact: Chris Town

TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TF
Class: sheared ends.
Location/Qualifiers
1. 739
/organism="Brassica oleracea"
/strain="TO1000DH3"
/db_xref="taxon:3712"
/clone="BOHRG31"

/clone.lib="BOHR"
/note="Vector: PHOS1; Site: 1; BstXI; 2-3 kb sheared
genomic DNA inserted into PHOS1 using BstXI linkers"
BASE COUNT 263 a 95 c 99 g 282 t
ORIGIN

Query Match 8.0%; Score 227; DB 12; Length 739;
Best Local Similarity 73.0%; Pred. No. 2.3e-15;
Matches 373; Conservative 0; Mismatches 105; Indels 33; Gaps 5;

Y 725 tctaatgctcaccacaaataaattcattatgcatatctatcttcttgaagtgta 784
DB 685 TCTAGATTCAACCACTAATAGATTATATTTCAATTCGATCTTTAAATGGA 626
Y 785 aacaaatattcaagttatattatgtttttaataaagaataaataaataa 844
DB 625 AACAAATATTAT-AGCTTATATTATGTTTAAATAGCAAGTAAACATAAATAA 567
Y 845 aataatagtagttacaaaaaataataattttaccagcgccanaaaaccta 904
DB 566 AATAGAGGAGTTACATATAATTAATTTTA-----AAAAATATT 526
Y 905 accaaacccaaatatttaacatttagttaaaccttaaccttggataatctaa 964
DB 525 TTTAAGCTGCTCATATAAACATTAAATCCTAAACCACTAAACCTTGAAGTTTAA 466
Y 965 attaaacattaaacacctaaccctaaccctaaccctaaccctaacccta 1024
DB 465 TCTAAATCAAAAATACTAAACACTAAACATTAACCTTAACCTTGAAGTTTAA 406
Y 1025 tttaagtttttgatttattagtttaggtttccaaaagtttaagtttaaccca 1084
DB 405 TTTAGTGTGTTGATTTAACCTTAGATTATCCAGGTTTATGTTTACC----- 351
Y 1085 ttatggttaggaattagacttagacttaggttaggttaggttaggttaggt 1143
DB 352 -AAGGTTTAGGATTAGGATTAGATTAGTGTGTTGCGTAGAGCGCTAAATAATTA 294
Y 1144 tttaaaaaatt 1199
DB 293 TTTAAAAAACATTTTGTGTAACCTACTACTATTTTATTATTATTATTATTAT 234
Y 1200 ttataataaaacataataataattatct 1230
DB 233 TTATTTTAAACATATAATTAACCTGAATA 203

RESULT 9
BH581638/c 614 bp DNA linear GSS 15-DEC-2001
LOCUS BOHNM45TR BOHL Brassica oleracea genomic clone BOHNM45, DNA
DEFINITION sequence.
ACCESSION BH581638
VERSION BH581638.1 GI:17833919
KEYWORDS GSS.
SOURCE Brassica oleracea.
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
1 (bases 1 to 614)
Town, C.D., Van Aken, S., Uterback, T. and Fraser, C.M.
Whole genome shotgun sequencing of Brassica oleracea
Unpublished (2001)
Other GSSs: BOHNM45TR
CONTACT: Chris Town
TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.

Seq primer: TF
FEATURES Class: sheared ends.
Location/Qualifiers
Source 1..614
/organism="Brassica oleracea"
/strain="TO1000DH3"
/db.xref="taxon:3712"
/clone="BOHNM45"
/clone.lib="BOHL"
/note="Vector: PHOS1; Site: 1; BstXI; 2-3 kb sheared
genomic DNA inserted into PHOS1 using BstXI linkers"
BASE COUNT 225 a 80 c 89 g 220 t
ORIGIN

Query Match 7.9%; Score 226.4; DB 12; Length 614;
Best Local Similarity 71.0%; Pred. No. 2.9e-15;
Matches 387; Conservative 0; Mismatches 117; Indels 41; Gaps 5;

Y 719 taaccccttaagttcaccacaaataaattcattatgcatatctatctttaga 778
DB 588 TGAACCTCTAGTGACCAACCACTATAGATTTCATATTTCAAATTTGATCTTTTAA 529
Y 779 aagtgaaacaaatattatcaagttatattatggtttttcaataaagaataaata 838
DB 528 ATAGGAACCAAAATATGTCAACTATATTACTTT-----AAAAAATAA 481
Y 839 ataaaaataatagtagttacaaaaaataataatttttccagcgccanaaac 898
DB 480 GTAAACCAAAATAGTATTACAGAAAATAATTAATTTT-----TCACCAAC 430
Y 899 actaaacccaaacccataataatttaacatttagttaaaccttaaccttggata 958
DB 429 ACTAAACCTTAATCCTATCCCTAACTCTGGGTAACCTTAACCTTGAATTAATC 370
Y 959 ttaa-----cattaaacattaaacactaaacccctaaccctaaccct 1003
DB 369 CTAACACTAAATAAAAAGCACTAACCACTAAACCTTAATTCCTAAACCT 310
Y 1004 aaaccccttaagttttaaattgtagttttttgttatagtttaggttaggttag 1063
DB 309 AAACCTTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTT 250
Y 1064 gtttaaggttagcccaagaagtttaggttaggttaggttaggttaggttaggt 1123
DB 249 GTTAGGGTTTATCCAGGTTTACGTTTGTATTTAGGTTTATGATTAATAGG 190
Y 1124 ctgacgacgttcaaaagttttttaaaaaatttttttttttttttttttt 1183
DB 189 GTTTAG-TGTTTGGCCGACGAGCTTAAATAATATTTTTGTATTTAGTA---TTTTT 135
Y 1184 ttatttttacccttttttttttttttttttttttttttttttttttttttt 1243
DB 134 ATTCTTTTACCTTTTAATTTTAAACATATAATTAATTTGACAAATATTTGTTCC 75
Y 1244 attaa 1248
DB 74 TTAA 70

RESULT 10
BH603601/c 790 bp DNA linear GSS 15-DEC-2001
LOCUS BOHLM68TR BOHL Brassica oleracea genomic clone BOHLM68, DNA
DEFINITION sequence.
ACCESSION BH603601
VERSION BH603601.1 GI:17856047
KEYWORDS GSS.
SOURCE Brassica oleracea.
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.

REFERENCE 1 (bases 1 to 790)
 AUTHORS Town, C.D., Van Aken, S., Uterback, T. and Fraser, C.M.
 TITLE Whole genome shotgun sequencing of Brassica oleracea
 JOURNAL Unpublished (2001)
 COMMENT Other-SSS: BOHLM68TF
 Contact: Chris Town

TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA.
 Tel: 301-838-3523
 Fax: 301-838-0208
 Email: cdtown@tigr.org
 DNA is from a doubled haploid provided by Tom Osborn.
 Seq primer: TR
 Class: sheared ends.

FEATURES

source 1. 790
 /organism="Brassica oleracea"
 /strain="TO1000DH3"
 /db_xref="taxon:3712"
 /clone="BOHLM68"
 /note="Vector: PHOS1; Site 1: BstXI; 2-3 kb sheared genomic DNA inserted into PHOS1 using BstXI linkers"
 BASE COUNT 299 a 123 c 91 g 277 t
 ORIGIN

Query Match 7.6%; Score 217.8; DB 12; Length 790;
 Best Local Similarity 69.6%; Pred. No. 2.1e-14;
 Matches 401; Conservative 0; Mismatches 153; Indels 22; Gaps 7;

OY 673 agtgaatcaccctcaacgtttgtaactatcttgcaccaccctcaaccttaagt 722
 DB 615 ATTATATATATATATATATATGAGAGTATTTTGGTTCATCCTAGAACCTAGT 556
 OY 733 caccacaataaattcaattatgcatactatcatttagaagtgaaacaaat 792
 DB 555 CACCAACCAATAGGATTTGTTATTTATTTGATATCTTCMAAAGTAACAAAT 496
 OY 793 attataaattatatttttttttttttttttttttttttttttttttttttt 852
 DB 495 ATTGTTAGTGTGTTATGTTTAAATATATATAGTAAATAAT-AAAAATATAG 437
 OY 853 tagttacaataaataaattt-----ataattttaccagcgtcanaaac 898
 DB 436 TAGTTACAAAAAGATTTTAAATAATTTATATTTTAACTGCTCAGCAAA 377
 OY 899 actaaacctaaacctaaatttaacttttagttaaacccttaacttgataatc 958
 DB 376 CACTAAACCAATTCCTAAACCTTAACCTTGCTAAACCCGAAACCTTGATTAATC 317
 OY 959 ttaaaatttaaacattaaacactaaaccttaacttaacttaacaccttaagt 1018
 DB 316 CTAAACACTTAACA-TCAAAACCTTAATCCTTAACCTTGAGTGTGTTAGT 258
 OY 1019 taaagttagttagtttttttttttttttttttttttttttttttttttttt 1077
 DB 257 TTGAGATTAGATTTATCA--AGGCTTAACATTATCCAAAGGTTTAGGTTAGCA 200
 OY 1078 -caagaatttagttagttagttagttagttagttagttagttagttagttagt 1136
 DB 199 TTGAGATTAGGATTAGGATTAGGATTAGGATTAGGATTAGGATTAGGATTAT 140
 OY 1137 aagatt 1194
 DB 139 ACTAAAGACGTTAAATATTTTGTGTAACACTACTATTTTATTTTATTTGTA 80
 OY 1195 ccttttatataaataataataataataataataataataataataataata 1230
 DB 79 CCTATTATTTTAAACATAGTATACCTTGACAAT 44

RESULT 11.

BH570528/c
 LOCUS BH570528 836 bp DNA linear GSS 14-DEC-2001
 DEFINITION BOG9808r BOG Brassica oleracea genomic clone BOG980, DNA
 sequence.
 ACCESSION BH570528
 VERSION BH570528.1 GI:17822367
 KEYWORDS GSS.
 SOURCE Brassica oleracea.
 ORGANISM Brassica oleracea.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
 REFERENCE 1 (bases 1 to 836)
 AUTHORS Town, C.D., Van Aken, S., Uterback, T. and Fraser, C.M.
 TITLE Whole genome shotgun sequencing of Brassica oleracea
 JOURNAL Unpublished (2001)
 COMMENT Other-SSS: BOG980TF
 Contact: Chris Town

TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA.
 Tel: 301-838-3523
 Fax: 301-838-0208
 Email: cdtown@tigr.org
 DNA is from a doubled haploid provided by Tom Osborn.
 Seq primer: TR
 Class: sheared ends.

FEATURES

source 1. 836
 /organism="Brassica oleracea"
 /strain="TO1000DH3"
 /db_xref="taxon:3712"
 /clone="BOG980"
 /note="Vector: PHOS1; Site 1: BstXI; 2-3 kb sheared genomic DNA inserted into PHOS1 using BstXI linkers"
 BASE COUNT 315 a 113 c 100 g 308 t
 ORIGIN

Query Match 7.2%; Score 206.8; DB 12; Length 836;
 Best Local Similarity 68.6%; Pred. No. 2.8e-13;
 Matches 385; Conservative 0; Mismatches 118; Indels 58; Gaps 5;

OY 736 caacacataaatttcaatttgcatacttctatctt-agaagtgaaacaaat 794
 DB 655 CACCCCTAATTTTCATTTTTCAAAATTCGATATCTTTTAAAGAAAGAAATAT 596
 OY 795 tatcaagttatattgatttttcaataaagaataaataaata-----aaaaa 848
 DB 595 TATCAAGTTATATATGTTTAAATAATTAATAATTAATAATTAATAATTAATA 536
 OY 849 atgtgtgttacaataaataaataataataattttaccaggtcamaaacaataaac 908
 DB 535 AAAGAAATTTTAAATAATATATTTTAAACATGCTCAGCAAAACCTTAATATCT 476
 OY 909 aaacccaataatttaacatttagttaaacccttaaccccttgataaatt----- 960
 DB 475 AATCCCTAATCCCTTAACCTTGCTTAACCTTAACCTTGCTTAACCTTAACCT 416
 OY 961 -----aaacattaaacattaaacactaaacactaaac 994
 DB 415 GATTAATCCTTAATATTAATTAATAATAAACACTAACCTTAACCTTAATATAT 356
 OY 995 ctaaaccttaaaccttaagtgtttaagtgtttagttagttagttagttagttagt 1054
 DB 355 ATTAACCTTAACCTTGATGTTTATGTTAGTATGATTTGATTTAGTATGATTT 296
 OY 1055 tatccaaagtttaagtttaaccagaagtttagt-----ttaaggataatgact 1107
 DB 295 TATCATAGAGTTTAGGTTTATCCAAAGGTTTAGGTTTAGGTTTAGGATTAGATT 236
 OY 1108 aggatttagttagttagttagttagttagttagttagttagttagttagttagt 1167

RESULT 12	LOCUS	DEFINITION
BH477804/c	BH477804	810 bp DNA
	BOHMS56GR BOHA	linear GSS 13-DEC-2001
	BOHMS56GR BOHA	plasmid clone BOHMS56, DNA
	sequence	

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 810)	Town, C.D., Van Aken, S., Utterback, T. and Fraser, C.M.	Whole genome shotgun sequencing of <i>Brassica oleracea</i>	Unpublished (2001)	Other GSSs: BOHAS56GFP

FEATURES	Location/Qualifiers
source	1. .810

```

location/Qualifiers
  1. 810
    /organism="Brassica oleracea"
    /strain="T01000DH3"
    /db_xref="taxon:3712"
    /clone="BOHASS6"
    /clone_1lb="BOHA"
    /note="Vector: pHOS1, site_1: BstXI; 2-3 kb sheared
genomic DNA inserted into pHOS1 using BstXI linkers
  280 a      95 c      123 g      312 t

```

Query Match	7.28;	Score 205.4;	DB 12;	Length 810;
Best Local Similarity	67.7%;	Prod. No. 4.1e-13;		
Matches 410;	Conservative 0;	Mismatches 127;	Indels 69;	Gaps 6

[illegible]

781 gtgaacacaatatattatcaagtatatatcgttttccaataaaaagat----- 829

830 -----AAAAAAAAATGCAAAAAATAGC 488
54/ AAGGACCAAAATATTATCAATTTATATGTTTTAAAAATAAAAAATAGC 488

[illegible]

881 ttaccagcgtcanaaacactaaaccctaaacc-----taataatctaact 928

929 ttagttaaacccctaacccttqataagaatctt-----aaagcatttccccctt- 368

RESULT	13
BH441036	
LOCUS	
DEFINITION	BH441036 810 bp DNA linear GSS 12-DEC-2001 BOHC R23TR BOHC Brassica oleracea genomic clone BOHC R23, DNA sequence

ACCESSION	BH441036
VERSION	BH441036.1
KEYWORDS	GI:17626750
SOURCE	GSS.
ORGANISM	Brassica oleracea.
	Brassica oleracea

Eukaryota: Viridiplantae; Streptophyta; Embryophyta: Tracheophyta
 Spermatophyta; Magnoliophyta; eudicotyledons: core eudicots:
 Rosidae: eurosids II: Brassicales; Brassicaceae; Brassica.
 1 (bases 1 to 810)
 Town, C.D., Van Aken, S., Uterback, T., and Fraser, C.M.
 Whole genome shotgun sequencing of *Brassica oleracea*
 unpublished (2001)
 Contact: Chris Town

9/12 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtrom@tigr.org

DNA is from a doubled haploid provided by Tom Osborn

Seq primer: TR

Class: sheared ends.

Class.	sheared ends.	Location/Qualifiers
FEATURES		
source		1. .810

```

/organism="Brassica oleracea"
/strain="T01000DH3"
/db_xref="taxon:3712"
/clone="BOHCR23"
/clone_1ib="BOHC"
/notes="Vector: PHOSI; Site_1: BstXI; 2-3 kb sheared
genomic DNA inserted into PHOSI using BstXI linkers
260 a 200 c 141 g 209 t
BASE COUNT
BRIGIN

```

Query Match	7.2%	Score 205;	DB 12;	Length 810;
Best Local Similarity	93.0%	Pred. No. 4.5e-13;		
Matches 226;	Conservative 0;	Mismatches 15;	Indels 2;	Gaps 1

QY 1 ggtctccacaaagaanaacgaagagaanaatggttcttcactcatataatatattgt 60
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 50 gggttcccacagaanaaccacagaagttaaggttccgcaccttcattaaatatatttttgt 100


```

OY 779 aagtgaaacaaatattatcaggtatattatgttttcaataaagaataaataa 838
Db 140 AAAGAAACAAATATGTGCAAGTATATATGTTTTTAAATAAAGGTAAATAA 199
OY 839 ataaaaataatagtagttacaataaataaataatatttaccaggtc---ana 894
Db 200 AATATAATATATATATGTTACAAATTTTTTTTAAATATATTTTTTAGGTCGTATCA 259
OY 895 aaacactaaacccaacccaataatttaacttttagtgaacccaacccaaccca 952
Db 260 AAACACTAAACCCCTAAATCCTAAACCAAGAAATCTTTGGTAAACCCCTAAATCCTAGCTC 319
OY 953 taaacttaaacattaaacattaaacccaacccaacccaacccaacccaaccca 1012
Db 320 TAAATATAAAACATTAACATTAACACCTAAACCCCTAAATCTTAAACCCCTAA 379
OY 1013 agtgtttaaatgttagtgttttggatctatagtttaggattatccaag 1063
Db 380 ACCCTGAGTGTGTTAGTGTGTTAGATTTAGATTTATGATGATGATCAAG 430

```

Search completed: June 28, 2002, 12:25:26
 Job time: 12151 sec